

GenCore version 5.1.1.8
Copyright (c) 1993 - 2006 Bioacceleration Ltd.
OM protein - protein search, using sw model
Run on: May 17, 2006, 10:55:27 ; Search time 201 Seconds
(without alignments)
646.017 Million cell updates/sec
Title: US-10-768-158-2
Perfect score: 1530
Sequence: 1 MAESEAETPSTPGFEFSKYF.....FDLVYKQKMGKDLTFDFYL 284
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 2589679 seqs, 457216429 residues
Total number of hits satisfying chosen parameters: 2589679
Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries
Database : A_Geneseq_8.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*
9: Geneseqp2005s.*
10: Geneseqp2006s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1530	100.0	284	3	AAV79215 Human tra
2	1530	100.0	284	5	AAU78268 Human ner
3	1530	100.0	284	6	ABR01809 Human can
4	1530	100.0	284	8	ADQ89158 Human uro
5	1530	100.0	284	8	ADR44884 Polypepti
6	1530	100.0	284	9	ADZ21933 Human SUL
7	1506	98.4	284	6	ABM04807 Rat sulfo
8	1002	65.5	218	6	ABR01860 Human can
9	867.5	56.7	171	6	ABR01861 Human can
10	867.5	56.7	171	9	ADZ21934 Human SUL
11	850.5	55.6	159	5	AAU78269 Rat nervo
12	480	31.4	296	5	ADII17131 Murine NO
13	474	31.0	1305	4	ABG27554 Novel hum
14	472	30.8	295	2	AAW23857 E6AP-bind
15	472	30.8	295	7	ADD18680 Human dis
16	472	30.8	295	8	ADP24302 PRO polyp
17	472	30.8	295	9	ADW71803 Human phe
18	467	30.5	296	5	ADII17130 Human can
19	458	29.9	285	8	ADY80433 Rat aryl
20	458	29.9	291	7	ADA45448 Rat Prote
21	458	29.9	291	7	ADE57139 Rat Prote
22	458	29.9	291	7	ADE57135 Rat Prote
23	458	29.9	291	7	ADE57147 Rat Prote

97 357.5 23.4 254 8 ABM82695 Abm82695 Human dia
 98 357.5 23.4 254 8 ABM82697 Abm82697 Human dia
 99 357.5 23.4 283 7 ADD48500 Add48500 Rat Prote
 100 356.5 23.3 285 10 AEF05774 Aef05774 Human hyd

ALIGNMENTS

RESULT 1

AAV79215

ID AAV79215 standard; protein; 284 AA.

XX

AC AAV79215;

XX

DT 19-JUN-2000 (first entry)

XX

DE Human transferase TRNSFS-7.

XX

KW Transferase; TRNSFS-7; human; antitumour; cancer;

KW gastrointestinal disorder; developmental disorder; genetic disorder;

KW neurological disorder; reproductive disorder; smooth muscle disorder;

KW immunological disorder; inflammation; diagnosis; therapy;

KW sulfotransferase.

XX

OS Homo sapiens.

XX

Key Location/Qualifiers

FH Modified-site 4

FT Modified-site /note= "potential O-phosphorylation"

FT Modified-site 11

FT Modified-site /note= "potential O-phosphorylation"

FT Domain 23..272

FT /note= "sulfotransferase protein domain"

FT Modified-site 104

FT Modified-site /note= "potential O-phosphorylation"

FT Modified-site 148

FT Modified-site /note= "potential O-phosphorylation"

FT Modified-site 153

FT Modified-site /note= "potential O-phosphorylation"

FT Modified-site 157

FT Modified-site /note= "potential O-phosphorylation"

FT Modified-site 172

FT Modified-site /note= "potential O-phosphorylation"

FT Modified-site 174

FT Modified-site /note= "potential O-phosphorylation"

FT Modified-site 205

FT Modified-site /note= "potential O-phosphorylation"

FT Modified-site 260

FT Modified-site /note= "potential O-phosphorylation"

XX

PN WO200014251-A2.

XX

PD 16-MAR-2000.

XX

PF 09-SEP-1999; 99WO-US020989.

XX

PR 10-SEP-1998; 98US-00150657.

XX

PR 04-NOV-1998; 98US-00186779.

XX

PR 11-MAY-1999; 99US-0133642P.

XX

PA (INCY-) INCYTE PHARM INC.

XX

PI Tang YT, Corley NC, Guegler KJ, Baughn MR, Lal P, Yue H;

PI Hillman JL, Azimzai Y;

XX

DR WPI: 2000-256996/22.

XX

DR N-PSDB; AA294207.

XX

PT Human transferase proteins useful for preventing, diagnosing and treating

PT cancers and developmental, gastrointestinal, genetic, immunological,

PT neurological, reproductive and smooth muscle disorders.

XX

PS Claim 1; Page 85-86; 113pp; English.

XX

CC The present sequence is that of human transferase TRNSFS-7, 1 of 15
 CC claimed novel human transferase proteins of the invention (see AAY79209-
 CC 23). The sequence was deduced from a cDNA clone (see AA294207) isolated
 CC from a cerebellum library. It shows homology to sulfotransferase
 CC proteins. TRNSFS-7 is expressed in nervous, endocrine and development
 CC tissues, especially those associated with neurological disorders,
 CC inflammation and cancer. The new human transferases and polynucleotides
 CC can be used in the diagnosis, prevention and treatment of cancer,
 CC developmental disorders, gastrointestinal disorders, genetic disorders,
 CC immunological disorders, neurological disorders, reproductive disorders,
 CC and smooth muscle disorders. The polypeptides can also be used to raise
 CC antibodies, and to screen for agonists and antagonists of transferase
 CC activity

XX

SQ Sequence 284 AA;

Query Match 100.0%; Score 1530; DB 3; Length 284;

Best Local Similarity 100.0%; Pred. No. 2.3e-151;

Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAESEAEPTSPGFEFESKYFFHGVRLPPFCRGKMEEIANFPVRPSDVMIYTPKSGTSL 60

Db 1 MAESEAEPTSPGFEFESKYFFHGVRLPPFCRGKMEEIANFPVRPSDVMIYTPKSGTSL 60

QY 61 LOEVVYLVSGADPDEIGLMNIDBOLPVLEYVPOQGLDIKELTSPRLIKSHLPVRFPSD 120

Db 61 LOEVVYLVSGADPDEIGLMNIDBOLPVLEYVPOQGLDIKELTSPRLIKSHLPVRFPSD 120

QY 121 LHNGDSKVIYMARNPKDLVSYQFPHRSRLRTWSYRGTFQFCRRPMNDKLGYSWFHVQ 180

Db 121 LHNGDSKVIYMARNPKDLVSYQFPHRSRLRTWSYRGTFQFCRRPMNDKLGYSWFHVQ 180

QY 181 EFWHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLALETCHQLVDPCCNA 240

Db 181 EFWHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLALETCHQLVDPCCNA 240

QY 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFFDYL 284

Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFFDYL 284

RESULT 2

AAU78268

ID AAU78268 standard; protein; 284 AA.

XX

AC AAU78268;

XX

DT 05-JUN-2002 (first entry)

XX

DE Human nervous system-specific sulfotransferase, SULTn, protein sequence.

XX

KW Human; nervous system-specific sulfotransferase; SULTn; sulfonation;

KW neurological disorder; multiple sclerosis; Huntington's chorea;

KW nervous system-mediated muscular atrophy; amyotrophic lateral sclerosis;

KW myasthenia gravis; Alzheimer's disease; senile dementia; Schizophrenia;

KW bipolar disorder; cytosolic sulfotransferase gene family;

KW cerebellar ataxia; chromosome 22q13.

XX

OS Homo sapiens.

XX

FH Key

FT Binding-site

FT Binding-site

FT Binding-site

FT Binding-site

FT Binding-site

FT Binding-site

XX

XX WO200218541-A2.

XX

XX 07-MAR-2002.

PD

Location/Qualifiers

54..61

/label= PAPS_5'_phosphosulfate_binding_site

129..144

/label= PAPS_3'_phosphate_binding_site

246..258

/label= PAPS_3'_phosphate_binding_site

XX

XX 04-SEP-2001; 2001WO-US041986.
XX 01-SEP-2000; 2000US-0229929P.
XX (UYBO-) UNIV BOSTON.
XX Farb DH, Martin S;
XX WPI; 2002-281063/32.
XX N-PSDB; ABK12289.
XX New nervous system-specific sulfotransferase (SULTn) polypeptide for
XX treating neurological disorders including multiple sclerosis, nervous
XX system-mediated muscular atrophy and Huntington's chorea.
XX Claim 2; Fig 1a; 46pp; English.
XX The present invention relates to a new nervous system-specific
XX sulfotransferase (SULTn) polypeptide. The SULTn polypeptide of the
XX invention comprises a nervous system-specific sulfotransferase,
XX comprising a sequence of 284 amino acids, given in the specification. The
XX invention is useful for identifying an effector of nervous system-
XX specific sulfonation by contacting the polypeptide with a candidate
XX effector compound and assaying the ability of the candidate effector
XX compound to effect the sulfotransferase function of the polypeptide. The
XX invention is also useful for inhibiting nervous system-specific
XX sulfonation by contacting the polypeptide under physiological conditions
XX with a compound capable of interacting with and inhibiting the
XX sulfonation activity of the polypeptide, where the compound is an
XX antibody, peptide, polypeptide, nucleic acid, organic molecule or an
XX inorganic molecule. The polypeptide of the invention is also useful for
XX treating a neurological disorder, including multiple sclerosis, nervous
XX system-mediated muscular atrophy, amyotrophic lateral sclerosis,
XX Huntington's chorea, myasthenia gravis, Alzheimer's disease, senile
XX dementia, schizophrenia, bipolar disorder and cerebellar ataxia. The
XX present amino acid sequence represents the human SULTn protein of the
XX invention. This sequence is encoded by the human SULTn gene located on
XX chromosome 22q13. SULTn is a member of the cytosolic sulfotransferase
XX gene family
SQ Sequence 284 AA;
Query Match 100.0%; Score 1530; DB 5; Length 284;
Best Local Similarity 100.0%; Pred. No. 2.3e-151;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MAESEATPSTPGEFESKYFEFHGVRLLPPFCRGKMEETANFVRPSDVWIVTPKSGTSL 60
Db 1 MAESEATPSTPGEFESKYFEFHGVRLLPPFCRGKMEETANFVRPSDVWIVTPKSGTSL 60
Qy 61 LQEVVYLVSOQADPDEIGLMNIDEQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LQEVVYLVSOQADPDEIGLMNIDEQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Qy 121 LHNGDSKVIYNARNPKDLVVSYYQFHRSLRTMSYRGTFQECRRFMNDKLGSGWFEHVQ 180
Db 121 LHNGDSKVIYNARNPKDLVVSYYQFHRSLRTMSYRGTFQECRRFMNDKLGSGWFEHVQ 180
Qy 181 EFWEHRMDSNVLFLKYEDMHRDLVTWVEQLARFLGVSCDKAQLAALTECHQLVDQCCNA 240
Db 181 EFWEHRMDSNVLFLKYEDMHRDLVTWVEQLARFLGVSCDKAQLAALTECHQLVDQCCNA 240
Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTTFDFYL 284
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTTFDFYL 284
RESULT 3
ID ABR01809
XX ABR01809 standard; protein; 284 AA.
XX AC ABR01809;

XX 19-MAY-2003 (first entry)
XX Human cancer-related protein, 192P2G7.
XX Human; cytostatic; vaccine; cancer; immune response.
XX Homo sapiens.
XX WO200283921-A2.
XX 24-OCT-2002.
XX 10-APR-2002; 2002WO-US011654.
XX 10-APR-2001; 2001US-0282739P.
XX 10-APR-2001; 2001US-0283112P.
XX 25-APR-2001; 2001US-0286630P.
XX (AGEN-) AGENSYS INC.
XX Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;
XX Morrison K, Morrison RK, Raitano AB;
XX WPI; 2003-075555/07.
XX N-PSDB; ABZ78140.
XX New composition comprising a substance that modulates the structure of
XX proteins and polynucleotides, useful for therapeutic, prognostic and
XX diagnostic reagents for eliciting cellular or humoral immune response in
XX cancer patients.
XX Claim 12; Fig 2R; 1021pp; English.
XX The present invention relates to novel human cancer-related genes and
XX proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
XX proteins are useful for eliciting a humoral or cellular immune response.
XX The genes are useful as probes and primers for the amplification and/or
XX detection of genes, mRNAs or their fragments, as reagents for the
XX diagnosis and/or prognosis of cancer, as coding sequences capable of
XX directing the expression of the protein, as tools for modulating or
XX inhibiting the expression of genes and/or translation of transcripts, and
XX as therapeutic agents. The proteins and peptides are useful as
XX therapeutic, prognostic and diagnostic reagents for cancer
SQ Sequence 284 AA;
Query Match 100.0%; Score 1530; DB 6; Length 284;
Best Local Similarity 100.0%; Pred. No. 2.3e-151;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MAESEATPSTPGEFESKYFEFHGVRLLPPFCRGKMEETANFVRPSDVWIVTPKSGTSL 60
Db 1 MAESEATPSTPGEFESKYFEFHGVRLLPPFCRGKMEETANFVRPSDVWIVTPKSGTSL 60
Qy 61 LQEVVYLVSOQADPDEIGLMNIDEQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LQEVVYLVSOQADPDEIGLMNIDEQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Qy 121 LHNGDSKVIYNARNPKDLVVSYYQFHRSLRTMSYRGTFQECRRFMNDKLGSGWFEHVQ 180
Db 121 LHNGDSKVIYNARNPKDLVVSYYQFHRSLRTMSYRGTFQECRRFMNDKLGSGWFEHVQ 180
Qy 181 EFWEHRMDSNVLFLKYEDMHRDLVTWVEQLARFLGVSCDKAQLAALTECHQLVDQCCNA 240
Db 181 EFWEHRMDSNVLFLKYEDMHRDLVTWVEQLARFLGVSCDKAQLAALTECHQLVDQCCNA 240
Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTTFDFYL 284
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTTFDFYL 284
RESULT 4

```

ADQ89158
ID ADQ89158 standard; protein; 284 AA.
XX
AC ADQ89158;
XX
DT 21-OCT-2004 (first entry)
XX
DE Human urological disorder related protein 16386 SEQ:110.
XX
KW urological disorder; uropathic; cytostatic; urinary incontinence;
KW benign prostatic hyperplasia; human.
XX
OS Homo sapiens.
XX
PN WO2004065576-A2.
XX
PD 05-AUG-2004.
XX
PF 14-JAN-2004; 2004WO-US000750.
XX
PR 15-JAN-2003; 2003US-0440318P.
PR 04-FEB-2003; 2003US-0444783P.
PR 27-MAR-2003; 2003US-0457901P.
PR 08-MAY-2003; 2003US-0468775P.
PR 19-MAY-2003; 2003US-0471614P.
PR 16-JUN-2003; 2003US-0478742P.
PR 18-JUL-2003; 2003US-0488529P.
PR 30-JUL-2003; 2003US-0491156P.
PR 02-SEP-2003; 2003US-049594P.
PR 26-SEP-2003; 2003US-0506332P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Karicheti V, Silos-Santiago I, Eliasof SD;
XX
DR WPI; 2004-562167/54.
XX
DR N-PSDB; ADQ89157.
XX
PT Use of polypeptides related to urological disorders, e.g. 44390, 54181,
PT 211 or for identifying a compound capable of treating a urological
PT disorder or identifying and treating a subject having a urological
PT disorder.
XX
PS Claim 1; SEQ ID NO 110; 542pp; English.
XX
CC The present invention describes the use of polypeptides related to
CC urological disorders for identifying a compound capable of treating a
CC urological disorder, identifying a subject having a urological disorder,
CC or treating a subject having a urological disorder. Also described: (1) a
CC method for identifying a compound capable of treating a urological
CC disorder; (2) a method for identifying a subject having a urological
CC disorder; and (3) a method for treating a subject having a urological
CC disorder. The compound has uropathic and cytostatic activities. The
CC polypeptides related to urological disorders are useful for identifying a
CC compound capable of treating a urological disorder, identifying a subject
CC having a urological disorder, or treating a subject having a urological
CC disorder. Disorders include urinary incontinence and benign prostatic
CC hyperplasia. The present sequence represents a human urological disorder
CC related protein, which is used in the exemplification of the present
CC invention.
XX
SQ Sequence 284 AA;
Query Match 100.0%; Score 1530; DB 8; Length 284;
Best Local Similarity 100.0%; Pred. No. 2.3e-151;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAESEAETPTGGEFESKYFEFHGVRLLPPFCRGKMEETANFVRPSDVIWVTPKSGTSL 60
DB 1 MAESEAETPTGGEFESKYFEFHGVRLLPPFCRGKMEETANFVRPSDVIWVTPKSGTSL 60
QY 61 LQEVVYLSQAGDPDEIGLNMNIDQLPVLEYQPQGLDIIKELTSRLIKSHLPYRFLPSD 120

```

```

Db 61 LQEVVYLSQAGDPDEIGLNMNIDQLPVLEYQPQGLDIIKELTSRLIKSHLPYRFLPSD 120
QY 121 LHNGDSKVIYMARNPKOLVVSYYQFHRSLRTMSVGRGTFOEFCRPFMDKLGYSWFEHVQ 180
Db 121 LHNGDSKVIYMARNPKOLVVSYYQFHRSLRTMSVGRGTFOEFCRPFMDKLGYSWFEHVQ 180
QY 181 EFWEHRMDSNVLFLLKYEDMHRDLVTMVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240
Db 181 EFWEHRMDSNVLFLLKYEDMHRDLVTMVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240
QY 241 EALPVGGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284
Db 241 EALPVGGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284

RESULT 5
ADR44884
ID ADR44884 standard; protein; 284 AA.
XX
AC ADR44884;
XX
DT 18-NOV-2004 (first entry)
XX
DE Polypeptide 16386 amino acid sequence.
XX
KW analgesic; gene therapy; pain; painful disorder.
XX
OS Homo sapiens.
XX
PN WO2004071411-A2.
XX
PD 26-AUG-2004.
XX
PF 30-JAN-2004; 2004WO-US002851.
XX
PR 04-FEB-2003; 2003US-0444781P.
PR 05-MAR-2003; 2003US-0452291P.
PR 13-MAR-2003; 2003US-0454540P.
PR 16-JUN-2003; 2003US-0478805P.
PR 30-JUL-2003; 2003US-0491048P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Silos-Santiago I, Karicheti V, Eliasof SD;
XX
DR WPI; 2004-625760/60.
DR N-PSDB; ADR44883.
XX
PT Identifying a compound capable of treating pain by combining a compound
PT to be tested with e.g., 16386 polypeptide or with a host cell expressing
PT the polypeptide and detecting the binding of the test compound to the
PT polypeptide.
XX
PS Disclosure; SEQ ID NO 2; 233pp; English.
XX
CC The invention relates to a method of identifying a compound capable of
CC treating pain or painful disorder by combining a compound to be tested
CC with an e.g., 16386, 15402, 21165, 1423, 636, or 32620 polypeptide or with
CC a host cell expressing the polypeptide and detecting the binding of the
CC test compound to the polypeptide to identify a compound that binds to the
CC polypeptide. The method is useful in identifying a compound capable of
CC treating pain or painful disorder. This sequence corresponds to the
CC polypeptide 16386 - a brain sulfoltransferase-like protein.
XX
SQ Sequence 284 AA;
Query Match 100.0%; Score 1530; DB 8; Length 284;
Best Local Similarity 100.0%; Pred. No. 2.3e-151;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAESEAETPTGGEFESKYFEFHGVRLLPPFCRGKMEETANFVRPSDVIWVTPKSGTSL 60
DB 1 MAESEAETPTGGEFESKYFEFHGVRLLPPFCRGKMEETANFVRPSDVIWVTPKSGTSL 60

```

Qy 61 LOEVVYLVSGADPDEIGLMNIDEQLPVLEYPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
|
Db 61 LOEVVYLVSGADPDEIGLMNIDEQLPVLEYPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
|
Qy 121 LHNGDSKVIYARNPKDLVVSYYQFHRSLRTMSYRGTFQECRRPMNDKLGYSWFHVQ 180
|
Db 121 LHNGDSKVIYARNPKDLVVSYYQFHRSLRTMSYRGTFQECRRPMNDKLGYSWFHVQ 180
|
Qy 181 EFWHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240
|
Db 181 EFWHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240
|
Qy 241 EALPVGRGRVGLWKDIFTVSNNEKFDLVYKQMGKCDLTFFPYL 284
|
Db 241 EALPVGRGRVGLWKDIFTVSNNEKFDLVYKQMGKCDLTFFPYL 284
|
RESULT 6
ADZ211933
ID ADZ211933 standard; protein; 284 AA.
XX
AC ADZ211933;
XX
DT 16-JUN-2005 (first entry)
XX
DE Human SULT4A1 splice variant 1 SEQ ID NO 1.
XX
KW neuroprotective; nootropic; diagnosis; prognosis; gene expression;
KW neurodegenerative disease; neurological disease; Alzheimers disease;
KW degeneration; sulfotransferase family 4A member 1; SULT4A1;
KW splice variant; enzyme.
XX
OS Homo sapiens.
XX
PN WO2005030947-A2.
XX
PD 07-APR-2005.
XX
PF 29-SEP-2004; 2004WO-EP052353.
XX
PR 30-SEP-2003; 2003US-0506775P.
XX
PA (EVOT-) EVOTEC NEUROSCIENCES GMBH.
XX
PI Von Der Kammer H, Pohlner J;
XX
XX WPI; 2005-262951/27.
DR N-PSDB; ADZ21938, ADZ21935.
XX
PT Diagnosing or prognosticating a neurodegenerative disease in a subject by
PT determining a level and/or activity of a transcription and/or translation
PT product of a gene coding for a cytosolic sulfotransferase family 4A
member 1.
XX
PS Claim 15; SEQ ID NO 1; 83pp; English.
XX
XX The invention describes a method of diagnosing or prognosticating a
CC neurodegenerative disease in a subject, or determining whether a subject
CC is at increased risk of developing the disease. The method comprises
CC determining a level and/or an activity of a transcription product and/or
CC a translation product of a gene coding for a cytosolic sulfotransferase
CC family 4A member 1 and/or a fragment, or derivative or variant of the
CC transcription or translation product, in a sample obtained from the
CC subject and comparing the level and/or the activity to a reference value
CC representing a known disease or health status, therefore diagnosing or
CC prognosticating the neurodegenerative disease in the subject, or
CC determining whether the subject is at increased risk of developing the
CC neurodegenerative disease. Protein molecules comprising fully defined 284
CC and/or 171 amino acid (SEQ ID NO. 1 and/or 2, respectively) sequence
CC given in the specification, the protein molecules being translation
CC products of the gene coding for a cytosolic sulfotransferase family 4A
CC member 1, or its fragments, or derivatives, or variants, are useful as

CC diagnostic targets for detecting a neurodegenerative disease, preferably
CC Alzheimer's disease or as screening targets for reagents or compounds
CC preventing, or treating, or ameliorating the disease. An antibody
CC specifically immunoreactive with an immunogen, where the immunogen is a
CC translation product of a gene coding for a cytosolic sulfotransferase
CC family 4A member 1, SEQ ID NO. 1 or 2, or its fragment, or derivative, or
CC variant is useful for detecting the pathological state of a cell in a
CC sample obtained from a subject, comprising immunocytochemical staining of
CC the cell with the antibody, where an altered degree of staining, or an
CC altered staining pattern in the cell compared to a cell representing a
CC known health status indicates a pathological state of the cell which
CC relates to a neurodegenerative disease, preferably to Alzheimer's
CC disease. The genetically altered non-human animal is useful for
CC screening, testing, and validating compounds, agents, and modulators in
CC the development of diagnostics and therapeutics to treat
CC neurodegenerative diseases, in particular Alzheimer's disease. An
CC agent(s) which directly or indirectly affect an activity and/or a level
CC of a gene coding for a cytosolic sulfotransferase family 4A member 1,
CC and/or a transcription product or a translation product of a gene coding
CC for a cytosolic sulfotransferase family 4A member 1, and/or their
CC fragment, or derivative, or variant is useful for treating or preventing
CC a neurodegenerative disease, in particular Alzheimer's disease. This is
CC the amino acid sequence of human sulfotransferase family 4A member 1
CC splice variant 1 (SULT4A1sv1).
XX
SQ Sequence 284 AA;
|
Query Match 100.0%; Score 1530; DB 9; Length 284;
Best Local Similarity 100.0%; Pred. No. 2.3e-151;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MAESEATPTSPGFEFSKYFEFHGVRLPPFCRGKMEBIANPPVRPSDVIIVTPKSGTSL 60
|
Db 1 MAESEATPTSPGFEFSKYFEFHGVRLPPFCRGKMEBIANPPVRPSDVIIVTPKSGTSL 60
|
Qy 61 LOEVVYLVSGADPDEIGLMNIDEQLPVLEYPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
|
Db 61 LOEVVYLVSGADPDEIGLMNIDEQLPVLEYPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
|
Qy 121 LHNGDSKVIYARNPKDLVVSYYQFHRSLRTMSYRGTFQECRRFMDKLGYSWFHVQ 180
|
Db 121 LHNGDSKVIYARNPKDLVVSYYQFHRSLRTMSYRGTFQECRRFMDKLGYSWFHVQ 180
|
Qy 181 EFWHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240
|
Db 181 EFWHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240
|
Qy 241 EALPVGRGRVGLWKDIFTVSNNEKFDLVYKQMGKCDLTFFPYL 284
|
Db 241 EALPVGRGRVGLWKDIFTVSNNEKFDLVYKQMGKCDLTFFPYL 284
|
RESULT 7
ABM04807
ID ABM04807 standard; protein; 284 AA.
XX
XX AC ABM04807;
XX
XX DT 22-SEP-2003 (first entry)
XX
XX DE Rat sulfotransferase-like protein.
XX
XX KW spinal cord; neuropathic pain; central sensitisation pain; pain;
XX analgesic; gene therapy.
XX
XX OS Rattus norvegicus.
XX
XX PN EPI284298-A2.
XX
XX PD 19-FEB-2003.
XX
XX PF 26-JUL-2002; 2002EP-00255229.
XX

PR	10-APR-2001; 2001US-0283112P.	PN	WO2005030947-A2.
PR	25-APR-2001; 2001US-0286630P.	XX	
XX		PD	07-APR-2005.
PA	(AGEN-) AGENSYS INC.	XX	
XX		PF	29-SEP-2004; 2004WO-EP052353.
XX	Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;	XX	
PI	Morrison K, Morrison RK, Raitano AB;	PR	30-SEP-2003; 2003US-0506775P.
XX		XX	(EVOT-) EVOTEC NEUROSCIENCES GMBH.
DR	WPI; 2003-075555/07.	PA	
DR	N-PSDB; ABZ78168.	XX	
XX		PI	Von Der Kammer H, Pohlner J;
XX	New composition comprising a substance that modulates the structure of	XX	
PT	proteins and polynucleotides, useful for therapeutic, prognostic and	DR	WPI; 2005-262951/27.
PT	diagnostic reagents for eliciting cellular or humoral immune response in	DR	N-PSDB; ADZ21936.
PT	cancer patients.	XX	
XX		PT	Diagnosing or prognosticating a neurodegenerative disease in a subject by
PS	Example 5; Fig 12r; 1021pp; English.	PT	determining a level and/or activity of a transcription and/or translation
XX		PT	product of a gene coding for a cytosolic sulfotransferase family 4A
CC		XX	member 1.
CC	The present invention relates to novel human cancer-related genes and	Claim 15; SEQ ID NO 2; 83pp; English.	
CC	proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and	XX	
CC	proteins are useful for eliciting a humoral or cellular immune response.	XX	
CC	The genes are useful as probes and primers for the amplification and/or	XX	
CC	detection of genes, mRNAs or their fragments, as reagents for the	CC	The invention describes a method of diagnosing or prognosticating a
CC	diagnosis and/or prognosis of cancer, as coding sequences capable of	CC	neurodegenerative disease in a subject, or determining whether a subject
CC	directing the expression of the protein, as tools for modulating or	CC	is at increased risk of developing the disease. The method comprises
CC	inhibiting the expression of genes and/or translation of transcripts, and	CC	determining a level and/or an activity of a transcription product and/or
CC	as therapeutic agents. The proteins and peptides are useful as	CC	a translation product of a gene coding for a cytosolic sulfotransferase
CC	therapeutic, prognostic and diagnostic reagents for cancer	CC	family 4A member 1 and/or a fragment, or derivative or variant of the
XX		CC	transcription or translation product, in a sample obtained from the
XX	Sequence 171 AA;	CC	subject and comparing the level and/or the activity to a reference value
Seq		CC	representing a known disease or health status, therefore diagnosing or
	Query Match 56.7%; Score 867.5; DB 6; Length 171;	CC	prognosticating the neurodegenerative disease in the subject, or
	Best Local Similarity 59.9%; Pred. No. 3.2e-82;	CC	determining whether the subject is at increased risk of developing the
	Matches 170; Conservative 1; Mismatches 0; Indels 113; Gaps 1;	CC	neurodegenerative disease. Protein molecules comprising fully defined 284
QY	1 MAESEAETPSTGPEFSKYFEFHGVRLPPFCRGMKEEIANFPVRPSDVWIVTPKSGTSL 60	CC	and/or 171 amino acid (SEQ ID NO. 1 and/or 2, respectively) sequence
Db		CC	given in the specification, the protein molecules being translation
	1 MAESEAETPSTGPEFSKYFEFHGVRLPPFCRGMKEEIANFPVRPSDVWIVTPKSGTSL 56	CC	products of the gene coding for a cytosolic sulfotransferase family 4A
QY	61 LQEVVVLVSGQADPDIGLWNIDQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRPLPSD 120	CC	member 1, or its fragments, or derivatives, or variants, are useful as
Db		CC	diagnostic targets for detecting a neurodegenerative disease, preferably
	57 -----	CC	Alzheimer's disease or as screening targets for reagents or compounds
QY	121 LHNGDSKVITYMARNPKDLVSVYQYFHRSLRTWSYRGTFQFCRRFMNDKLGYSWFHVQ 180	CC	preventing, or treating, or ameliorating the disease. An antibody
Db		CC	specifically immunoreactive with an immunogen, where the immunogen is a
	57 -----VGYGSWFHVQ 67	CC	translation product of a gene coding for a cytosolic sulfotransferase
QY	181 EFWEHRMDSNVLFKYEDMHRDLVTWVQIARPLGVSCDKAQLTEALTECHQLVDQCCNA 240	CC	family 4A member 1, SEQ ID NO. 1 or 2, or its fragment, or derivative, or
Db		CC	variant is useful for detecting the pathological state of a cell in a
	68 EFWEHRMDSNVLFKYEDMHRDLVTWVQIARPLGVSCDKAQLTEALTECHQLVDQCCNA 127	CC	sample obtained from a subject, comprising immunocytochemical staining of
QY	241 EALPVGRGVLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFVL 284	CC	the cell with the antibody, where an altered degree of staining or an
Db		CC	altered staining pattern in the cell compared to a cell representing a
	128 EALPVGRGVLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFFL 171	CC	known health status indicates a pathological state of the cell which
	RESULT 10	CC	relates to a neurodegenerative disease, preferably to Alzheimer's
ID	ADZ21934	CC	disease. The genetically altered non-human animal is useful for
XX	ADZ21934 standard; protein; 171 AA.	CC	screening, testing, and validating compounds, agents, and modulators in
AC	ADZ21934;	CC	the development of diagnostics and therapeutics to treat
DT	16-JUN-2005 (first entry)	CC	neurodegenerative diseases, in particular Alzheimer's disease. An
XX		CC	agent(s) which directly or indirectly affect an activity and/or a level
DE	Human SULT4A1 splice variant 2 SEQ ID NO 2.	CC	of a gene coding for a cytosolic sulfotransferase family 4A member 1,
XX		CC	and/or a transcription product or a translation product of a gene coding
KW	neuroprotective; nootropic; diagnosis; prognosis; gene expression;	CC	for a cytosolic sulfotransferase family 4A member 1, and/or their
KW	neurodegenerative disease; neurological disease; Alzheimer's disease;	CC	fragment, or derivative, or variant is useful for treating or preventing
KW	degeneration; sulfotransferase family 4A member 1; SULT4A1;	CC	a neurodegenerative disease, in particular Alzheimer's disease. This is
XX	splice variant; enzyme.	CC	the amino acid sequence of human sulfotransferase family 4A member 1
OS	Homo sapiens.	CC	splice variant 2 (SULT4A1svl).
XX		XX	
		Seq	Sequence 171 AA;
			Query Match 56.7%; Score 867.5; DB 9; Length 171;
			Best Local Similarity 59.9%; Pred. No. 3.2e-82;
			Matches 170; Conservative 1; Mismatches 0; Indels 113; Gaps 1;
QY	1 MAESEAETPSTGPEFSKYFEFHGVRLPPFCRGMKEEIANFPVRPSDVWIVTPKSGTSL 60		
Db			
	1 MAESEAETPSTGPEFSKYFEFHGVRLPPFCRGMKEEIANFPVRPSDVWIVTPKSGTSL 56		
	61 LQEVVVLVSGQADPDIGLWNIDQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRPLPSD 120		
	57 -----		
	121 LHNGDSKVITYMARNPKDLVSVYQYFHRSLRTWSYRGTFQFCRRFMNDKLGYSWFHVQ 180		
	57 -----VGYGSWFHVQ 67		
	181 EFWEHRMDSNVLFKYEDMHRDLVTWVQIARPLGVSCDKAQLTEALTECHQLVDQCCNA 240		
	68 EFWEHRMDSNVLFKYEDMHRDLVTWVQIARPLGVSCDKAQLTEALTECHQLVDQCCNA 127		
	241 EALPVGRGVLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFVL 284		
	128 EALPVGRGVLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFFL 171		
	RESULT 10		
ID	ADZ21934		
XX	ADZ21934 standard; protein; 171 AA.		
AC	ADZ21934;		
DT	16-JUN-2005 (first entry)		
XX			
DE	Human SULT4A1 splice variant 2 SEQ ID NO 2.		
XX			
KW	neuroprotective; nootropic; diagnosis; prognosis; gene expression;		
KW	neurodegenerative disease; neurological disease; Alzheimer's disease;		
KW	degeneration; sulfotransferase family 4A member 1; SULT4A1;		
XX	splice variant; enzyme.		
OS	Homo sapiens.		
XX			

QY 61 LQEVVYLVSQGADPDEIGLWNIDEQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120
 Db 57 ----- 56
 QY 121 LHNGSKVIYMARNP KD LVVSYQPHRS LRTMSYRGTFQEFRCRPMNDKLGYSWFHVQ 180
 Db 57 -----VGYGSWFHVQ 67
 QY 181 EFWEHRMDSNVLFKYEDMHRDLVTMVQLARFLGVSCDKAQLAEALTEHCHQLVDQCCNA 240
 Db 68 EFWEHRMDSNVLFKYEDMHRDLVTMVQLARFLGVSCDKAQLAEALTEHCHQLVDQCCNA 127
 QY 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFFDYL 284
 Db 128 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFFDYL 171

RESULT 11
 AAU78269
 ID AAU78269 standard; protein; 159 AA.
 XX
 AC AAU78269;
 DT 05-JUN-2002 (first entry)
 XX
 DE Rat nervous system-specific sulfotransferase, SULTn, protein sequence.
 XX
 KW Rat; nervous system-specific sulfotransferase; SULTn; sulfonation;
 KW neurological disorder; multiple sclerosis; Huntington's chorea;
 KW nervous system-mediated muscular atrophy; amyotrophic lateral sclerosis;
 KW myasthenia gravis; Alzheimer's disease; senile dementia; Schizophrenia;
 KW bipolar disorder; cytosolic sulfotransferase gene family;
 KW cerebellar ataxia.
 XX
 OS Rattus sp.
 PN WO200218541-A2.
 XX
 PD 07-MAR-2002.
 XX
 PF 04-SEP-2001; 2001WO-US041986.
 XX
 PR 01-SEP-2000; 2000US-0229929P.
 XX
 PA (UYBO-) UNIV BOSTON.
 XX
 PI Farb DH, Martin S;
 XX
 DR WPI; 2002-281063/32.
 XX
 DR P-PSDB; AAU78269.
 XX
 PT New nervous system-specific sulfotransferase (SULTn) polypeptide for
 PT treating neurological disorders including multiple sclerosis, nervous
 PT system-mediated muscular atrophy and Huntington's chorea.
 XX
 PS Claim 2; Page; 46pp; English.
 XX
 CC The present invention relates to a new nervous system-specific
 CC sulfotransferase (SULTn) polypeptide. The SULTn polypeptide of the
 CC invention comprises a nervous system-specific sulfotransferase,
 CC comprising a sequence of 284 amino acids, given in the specification. The
 CC invention is useful for identifying an effector of nervous system-
 CC specific sulfonation by contacting the polypeptide with a candidate
 CC effector compound and assaying the ability of the candidate effector
 CC compound to affect the sulfotransferase function of the polypeptide. The
 CC invention is also useful for inhibiting nervous system-specific
 CC sulfonation by contacting the polypeptide under physiological conditions
 CC with a compound capable of interacting with and inhibiting the
 CC sulfonation activity of the polypeptide, where the compound is an
 CC antibody, peptide, polypeptide, nucleic acid, organic molecule or an
 CC inorganic molecule. The polypeptide of the invention is also useful for
 CC treating a neurological disorder, including multiple sclerosis, nervous
 CC system-mediated muscular atrophy, amyotrophic lateral sclerosis,

CC Huntington's chorea, myasthenia gravis, Alzheimer's disease, senile
 CC dementia, Schizophrenia, bipolar disorder and cerebellar ataxia. The
 CC present amino acid sequence represents the rat SULTn protein of the
 CC invention. SULTn is a member of the cytosolic sulfotransferase gene
 CC family. Note: The sequence data for this sequence did not form part of
 CC the printed specification but was obtained in electronic format from
 CC GENBANK AF176343
 XX
 SQ Sequence 159 AA;
 Query Match 55.6%; Score 850.5; DB 5; Length 159;
 Best Local Similarity 98.1%; Pred. No. 1.7e-80;
 Matches 157; Conservative 2; Mismatches 0; Indels 1; Gaps 1;
 QY 125 DSKVIYMARNP KD LVVSYQPHRS LRTMSYRGTFQEFRCRPMNDKLGYSWFHVQEFWE 184
 Db 1 DSKVIYMARNP KD LVVSYQPHRS LRTMSYRGTFQEFRCRPMNDKLGYSWFHVQEFWE 60
 QY 185 HRMDSNVLFKYEDMHRDLVTMVQLARFLGVSCDKAQLAEALTEHCHQLVDQCCNAEALP 244
 Db 61 HRMDANVLFKYEDMHRDLVTMVQLARFLGVSCDKAQLAEALTEHCHQLVDQCCNAEALP 119
 QY 245 VGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFFDYL 284
 Db 120 VGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFFDYL 159

RESULT 12
 ADI17131
 ID ADI17131 standard; protein; 296 AA.
 XX
 AC ADI17131;
 XX
 DT 15-APR-2004 (first entry)
 XX
 DE Murine NOVX protein homologue SeqID 667.
 XX
 KW mouse; murine; NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes;
 KW inflammation; autoimmune disorder; allergy; blood disorder;
 KW acquired immunodeficiency syndrome; AIDS; obesity; asthma;
 KW immunoglobulin (Ig) A nephropathy; cirrhosis; arthritis;
 KW Alzheimer's disease; infect.
 XX
 OS Mus musculus.
 XX
 PN WO200268649-A2.
 XX
 PD 06-SEP-2002.
 XX
 PF 31-JAN-2002; 2002WO-US002785.
 XX
 PR 31-JAN-2001; 2001US-0265395P.
 PR 31-JAN-2001; 2001US-0265412P.
 PR 31-JAN-2001; 2001US-0265514P.
 PR 31-JAN-2001; 2001US-0265517P.
 PR 02-FEB-2001; 2001US-0266406P.
 PR 05-FEB-2001; 2001US-0266767P.
 PR 07-FEB-2001; 2001US-0266975P.
 PR 07-FEB-2001; 2001US-0267057P.
 PR 08-FEB-2001; 2001US-0267459P.
 PR 09-FEB-2001; 2001US-0267823P.
 PR 15-FEB-2001; 2001US-0268974P.
 PR 26-FEB-2001; 2001US-0271664P.
 PR 27-FEB-2001; 2001US-0271839P.
 PR 27-FEB-2001; 2001US-0271855P.
 PR 02-MAR-2001; 2001US-0272788P.
 PR 02-MAR-2001; 2001US-0273046P.
 PR 14-MAR-2001; 2001US-0275925P.
 PR 14-MAR-2001; 2001US-0275947P.
 PR 14-MAR-2001; 2001US-0275950P.
 PR 14-MAR-2001; 2001US-0275989P.
 PR 15-MAR-2001; 2001US-0276448P.
 PR 15-MAR-2001; 2001US-0276450P.

CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostic, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 1305 AA;

Query Match 31.0%; Score 474; DB 4; Length 1305;
 Best Local Similarity 33.5%; Pred. No. 1e-39; Mismatches 54; Gaps 8;
 Matches 110; Conservative 54; Indels 54; Gaps 8;

QY 3 ESEATPTSP-----GEFESKYPFEHGVRL-----PP--FCRG-----K 34
 DB 265 EGRQPTPTPTILSPCRQGENSGSRELNMELIQTSPPLEYKGVPLIKYFAEA 324
 QY 35 MEEIANFPVRPSDVIVTPKSGTSLLOEVVYLVSGQADPDIGLNMIDQLPVLEYPQP 94
 DB 325 LGPLOSFOARPDDLINTYPKSGTTWVSQILDIMYQGGDLEKCNRAPIYVRVPFLEVNDP 384
 QY 95 ---GLDIIKELTSRLIKSHLPYRFLPSDLHNGDSKVIYARNPKDLVVSYYQFHSRLR 150
 DB 385 GPSPGLETLKDTPPRLIKSHLPALLPQTLDDQKVVVYVARNPKDVAIVYHPRMEK 444
 QY 151 TMSYRGTFQFCRRFMNDKLGYSWFEHVOEFWEHRMDSNVLFKLYEDMHRDLVTMVQOL 210
 DB 445 AHPEGTWDSFLEKFMAGEVSGSYQHVQEWELSRTHPVLYLYFYEDMKNPKREIQKI 504
 QY 211 ARFLGVSCDKAQLALTEHC-----HQLVQCCNBAALP-VGRGRVGL 252
 DB 505 LEFVGRSLPEETDMFVQHTSPKTKKNPMTNTYTPQELMDHSIS----PPMRKGMAGD 560
 QY 253 WKDIETVSMNEKFDLVYKQKGCCLTF 280
 DB 561 WKTTFVAQNERFDADYAEKMGAGCSLSF 588

RESULT 14
 AAW23657
 ID AAW23657 standard; protein; 295 AA.
 XX
 AC AAW23657;
 XX
 DT 10-OCT-1997 (first entry)
 XX
 DE E6AP-binding protein cln25.
 XX
 KW Human; cell differentiation; survival; carcinoma; psoriasis;
 KW ubiquitination; p53; tumour suppressor; homeostasis; papilloma virus;
 KW epithelial cell; acne; ichthyosis; aphthous ulcer; hair growth; antibody;
 KW cell proliferation.
 XX
 OS Homo sapiens.
 XX
 PN WO9640767-A2.
 XX
 PD 19-DEC-1996.
 XX
 XX 06-JUN-1996; 96WO-US009040.
 PF
 XX 07-JUN-1995; 95US-00484878.
 PR
 XX

PA (MITO-) MITOTIX INC.
 XX Beer-Romero PL, Draetta G, Rolfe M;
 XX WPI; 1997-087053/08.
 DR N-PSDB; AAT78309.
 XX
 PT E6AP-binding proteins and related nucleic acid - useful for modulating
 PT cell differentiation, survival etc., partic. for treatment and diagnosis
 PT of carcinoma, psoriasis, etc.
 XX
 PS Claim 1; Page 70-71; 83pp; English.
 XX
 CC The present sequence represents the E6AP-binding protein cln25. E6AP
 CC mediates ubiquitination and so the inactivation of e.g. p53 tumour
 CC suppressor. The E6AP-binding protein is likely to be involved in normal
 CC cell homeostasis and in the pathogenesis of proliferative and
 CC differentiation disorders, e.g. regulation of gene expression or the cell
 CC cycle, modification of cell surface receptors, biogenesis of ribosomes
 CC and DNA repair. The protein, which can optionally be generated in vivo by
 CC gene therapy, may be useful in treatment and prevention of papilloma virus
 CC infected transformed cells and carcinoma, and may also be used to
 CC regulate epithelial cell processes more generally, e.g. in cases of
 CC psoriasis, acne, ichthyosis, aphthous ulcers. It may also be used to
 CC control wound healing, and inhibit growth of hair. It can also be used to
 CC generate antibodies which are used in immunoassays to determine the
 CC protein levels. The encoding nucleic acid can be used to prepare
 CC recombinant proteins and oligonucleotides useful as probes and primers
 CC for diagnostic detection of (mutant) mRNA for the protein in
 CC (transformed) cells, also for antisense therapy. It can also be used for
 CC the detection of mutations in E6AP-binding protein-encoding genes, mis-
 CC expression of these genes or mis-incorporation of the protein in a
 CC transcription regulatory complex containing E6AP can be used to assess
 CC risk of disorders characterised by cell proliferation
 XX
 SQ Sequence 295 AA;

Query Match 30.8%; Score 472; DB 2; Length 295;
 Best Local Similarity 35.8%; Pred. No. 2e-40;
 Matches 100; Conservative 52; Mismatches 101; Indels 26; Gaps 4;

QY 24 GVRPPFCRCGRKMERIANFPVRPSDVIVTPKSGTSLLOEVVYLVSGQADPDEIGLNMID 83
 DB 17 GVPLIKYFAEALGPLQSFQARPDDLINTYPKSGTTWVSQILDIMYQGGDLEKCNRAPIY 76
 QY 84 BOLPVLEYPQP-----GLDIIKELTSRLIKSHLPYRFLPSDLHNGDSKVIYARNPKDLV 139
 DB 77 VRVPFLEVNDPEFSGLETLKDTPPRLIKSHLPALLPQTLDDQKVVVYVARNPKDVA 136
 QY 140 VSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVOEFWEHRMDSNVLFKLYEDM 199
 DB 137 VSYHFRMEKAHPDPGTWDSFLEKFMAGEVSGSYQHVQEWELSRTHPVLYLYFYEDM 196
 QY 200 HRDLVTMVQELARFLGVSCDKAQLALTEHC-----HQLVQCCNBAE 242
 DB 197 KENPKREIQKILEFVGRSLPEETDMFVQHTSPKTKKNPMTNTYTPQELMDHSIS--- 253
 QY 243 LP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKGCCLTF 280
 DB 254 -PPMRKGMAGDKTTFVAQNERFDADYAEKMGAGCSLSF 291

RESULT 15
 ADD18680
 ID ADD18680 standard; protein; 295 AA.
 XX
 AC ADD18680;
 XX
 DT 15-JAN-2004 (first entry)
 XX
 DE Human disease related protein SeqID111.
 XX
 KW human; disease state; cytostatic; antiinflammatory; ophthalmological;

KW antiarteriosclerotic; vulnery; gene therapy;
KW hypoxia-regulated condition; tumorigenesis; angiogenesis; apoptosis;
KW inflammation; erythropoiesis; glycolysis; gluconeogenesis;
KW glucose transportation; catecholamine synthesis; iron transport;
KW nitric oxide synthesis; cancer; ischaemic condition; reperfusion injury;
KW retinopathy; neonatal stress; pre-eclampsia; atherosclerosis;
KW inflammatory condition; wound healing.
XX
OS Homo sapiens.
XX
XX WO2003018621-A2.
XX
XX PD 06-MAR-2003.
XX
XX PF 23-AUG-2002; 2002WO-GB003892.
XX
XX PR 23-AUG-2001; 2001GB-00020558.
XX
XX PR 05-OCT-2001; 2001GB-00024037.
XX
XX PA (OXFO-) OXFORD BIOMEDICA UK LTD.
XX
XX PI Kingsman SM, White J, Ward NR, Harris RA, Naylor S, Mundy CR;
XX
XX DR WPI: 2003-290046/28.
XX
XX DR N-PSDB; ADD18681.
XX
XX PT New substantially purified polypeptide, useful for diagnosing or treating
XX a hypoxia-regulated condition, such as cancer, ischemia, reperfusion
XX injury, retinopathy, pre-eclampsia, atherosclerosis, inflammation, or
XX wound healing.
XX
XX PS Claim 25; SEQ ID NO 111; 424pp; English.
XX
XX CC This invention relates to novel human genes and gene product which are
XX implicated in certain disease states. Compounds which modulate the
XX proteins of the invention may have cytostatic, antiinflammatory,
XX ophthalmological, antiarteriosclerotic or vulnary activities. The
XX sequences of the invention may be useful for gene therapy. The invention
XX may be useful for diagnosing or treating a hypoxia-regulated condition,
XX such as tumorigenesis, angiogenesis, apoptosis, inflammation,
XX erythropoiesis, or the biological response to hypoxia conditions
XX including processes such as glycolysis, gluconeogenesis, glucose
XX transportation, catecholamine synthesis, iron transport or nitric oxide
XX synthesis. The disease includes cancer, ischaemic conditions, reperfusion
XX injury, retinopathy, neonatal stress, pre-eclampsia, atherosclerosis,
XX inflammatory conditions or wound healing. The present sequence is that of
XX a disease related protein of the invention.
XX
XX SQ Sequence 295 AA;
Query Match 30.8%; Score 472; DB 7; Length 295;
Best Local Similarity 35.8%; Pred. No. 2e-40;
Matches 100; Conservative 52; Mismatches 101; Indels 26; Gaps 4;
QY 24 GVRPPFCRGKMEETANFVRPSDVWIVTPKSGTSLQEVYLVSQGADPEIGLMNID 83
DB ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
17 GVPLIKYFAEALGPLQSFQARPDLLINTYPKSGTTWVSQILDMYQGDLEKCNRAPIY 76
QY ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
84 EQLPVLVEYQPQ----GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYVARNPKDLV 139
DB ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
77 VRVPFLEVNDGPEPSGLTLDKTPPRLIKSHLPALLPQTLDDQKVKVYVARNPKDVA 136
QY ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
140 VSYQFHRSLRTMSYRGTFQBFRCFRFMNDKLGYSWFHFVQEFWHRMDSNVFLFKYEDM 199
QY ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
137 VSYVHFHMEKXAHPEPGTWDQSFLEKFMAGEVSYGSWYQVHVEWELSTHPLVLYLFYEDM 196
DB ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 200 HRDLVTWVEQLARFLGVSCDKAQLAEALTEHC-----HQLVDQCCNAEA 242
DB ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
197 KENPKREIQKILEFVGRSLPSETMDFMVQHTSFKEMKKNPMNTYTVTPQELMDHSIS--- 253
QY ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
243 LP-VGRGRVGLWKDIFTVSMNEKEDLVYKQMGKCDLTF 280
DB ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
254 -PFMRKGMAGDWKTTFTTVAQNERFDADYAEKMGAGCSLSF 291

RESULT 16
ADP24302
ID ADP24302 standard; protein; 295 AA.
XX
XX AC ADP24302;
XX
XX DT 18-NOV-2004 (first entry)
XX
XX DE PRO polypeptide SEQ ID NO:1480.
XX
XX KW PRO; antiinflammatory; antiarthritic; antirheumatic; immunosuppressive;
KW osteopathic; antidiabetic; dermatological; antipsozotic; antiallergic;
KW antiasthmatic; hepatotropic; respiratory; gene therapy; immune system.
XX
XX OS Unidentified.
XX
XX PN WO2004041170-A2.
XX
XX PD 21-MAY-2004.
XX
XX PF 30-OCT-2003; 2003WO-US034312.
XX
XX PR 01-NOV-2002; 2002US-0423394P.
XX
XX PA (GETH) GENENTECH INC.
XX
XX PI Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI;
PI Wu TD;
XX
XX DR WPI: 2004-419628/39.
XX
XX DR N-PSDB; ADP24301.
XX
XX PT New PRO polypeptides and polynucleotides, useful for treating e.g.
PT erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated
PT renal disease, or demyelinating diseases of the central or peripheral
XX nervous system.
XX
XX PS Claim 7; SEQ ID NO 1480; 2940pp; English.
XX
XX CC The invention relates to a novel isolated nucleic acid and the PRO
XX polypeptide encoded by it. A protein of the invention has
XX antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,
XX osteopathic, antidiabetic, dermatological, antipsozotic, antiallergic,
XX antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide
XX of the invention may have a use in gene therapy. The PRO polypeptide, its
XX agonist, antagonist, or antibody that specifically binds to the
XX polypeptide is useful for treating an immune related disorder such as
XX systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,
XX juvenile chronic arthritis, a spondyloarthopathy, systemic sclerosis, an
XX idiopathic inflammatory myopathy, Sjogren's syndrome, systemic
XX vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune
XX thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal
XX disease, a demyelinating disease of the central or peripheral nervous
XX system, idiopathic demyelinating polynuropathy, Guillain-Barre syndrome,
XX a chronic inflammatory demyelinating polynuropathy, a hepatobiliary
XX disease, infectious or autoimmune chronic active hepatitis, primary
XX biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,
XX inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's
XX disease, an autoimmune or immune-mediated skin disease, a bullous skin
XX disease, erythema multiforme, contact dermatitis, psoriasis, an allergic
XX disease, asthma, allergic rhinitis, atopic dermatitis, food
XX hypersensitivity, urticaria, an immunologic disease of the lung,
XX eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity
XX pneumonitis, a transplantation associated disease, graft rejection or
XX graft-versus-host disease. The present sequence represents a PRO protein
XX of the invention.
XX
XX SQ Sequence 295 AA;
Query Match 30.8%; Score 472; DB 8; Length 295;
Best Local Similarity 35.8%; Pred. No. 2e-40;

Matches	100;	Conservative	52;	Mismatches	101;	Indels	26;	Gaps	4
Qy	24	GVRLPFCRCGMKIEANFPVPRPSDVWIVTYPKSGTSLLEQEVVLVSQGADPEIGLMNID	83						
Db	17	GVPLIKYFAEALGFLSQFARPDLLINTYPKSGTTWVSQILDMIYQGGDLEKCNRAPIY	76						
Qy	84	EQLPVLEYQPQ-----GLDIHKELTSPLRIKSHLPHYRFLPSPDLHNGDSKVIIYARNPKDLV	139						
Db	77	VRVPFLEVDPGEPSGLETLXDTPPRLIKSHLPLALLPQTLLDOKVKVVYVARNPKDVA	136						
Qy	140	VSYIQFHRSLSRTMSYRGTFQEFRCRFMNDKLGYGSWFHFVOEFWEHRMDSNVLFLLKYEDM	199						
Db	137	VSYHYFHRMEKAHPPEPGTWDSFLFKFMAGEVSYGSWYQHVOEWBELSTHTPHVLILFYEDM	196						
Qy	200	HRDLVTWVEQLARTLGLVSCDKAQLALETHC-----HQLVDQCNAEA	242						
Db	197	KENPKRIQKILEFVGRLSPLPETWDPMVQHTSFKEMKKNPMNTYTVQELMDHSIS---	253						
Qy	243	LP-VGRGRVGLWKDIFTVMNEKFDLVLYKQMGKCDLTF	280						
Db	254	-PFMRKGMAGDWKKTFTTVAQNERFDADYAENKMGCSLSF	291						
RESULT	17								
ADW71803	ID	ADW71803 standard; protein; 295 AA.							
XX	AC	ADW71803;							
XX	DT	07-APR-2005 (first entry)							
XX	DE	Human phenol sulfotransferase protein, SULTIA3, SEQ ID 4.							
XX	KW	Selectable marker; screening; animal breeding; polymorphism;							
KW	KW	restriction fragment length polymorphism; allelic variation;							
KW	KW	aryl sulfotransferase; PST; phenol sulfotransferase; SULTIA3; enzyme.							
OS	OS	Homo sapiens.							
XX	FH	Key Location/Qualifiers							
FT	Active-site	83							
FT	Active-site	134							
FT	Active-site	263							
XX	PB	US2005019788-A1.							
XX	PD	27-JAN-2005.							
XX	PF	30-JAN-2004; 2004US-00769507.							
XX	PR	08-APR-1998; 98US-0081037P.							
PR	PR	08-APR-1999; 99US-00288037.							
XX	PR	23-NOV-2001; 2001US-00024628.							
XX	PA	(UYGU-) UNIV GUELPH.							
XX	FI	Squires EJ, Lin Z, Lou Y;							
XX	DR	WPI; 2005-111909/12.							
XX	PT	Genetically typing animals to determine those with desired boar taint characteristics, comprises obtaining a sample of genetic material from the animal, and assaying for the presence of a sulfotransferase allele.							
XX	PS	Example; SEQ ID NO 4; 24pp; English.							
XX	CC	The present invention relates to a method of genetically typing animals to determine those with desired boar taint characteristics. The method involves obtaining a sample of genetic material from the animal and assaying for the presence of a sulfotransferase allele. The present sequence is the human phenol sulfotransferase (PST) protein, SULTIA3. SULTIA3 protein encoding gene is located on chromosome 16p12.1.							


```
XX (GHO ) GEN HOSPITAL CORP.
PA (FARB ) BAYER AG.
XX
XX Woolf C, D'urso D, Befort K, Costigan M;
XX
XX WPI; 2003-268312/26.
DR GENBANK; P17988.
XX
XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
XX Claim 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regularly
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for producing a compound useful in treating
CC pain (e.g. spinal segmental nerve injury (SNI)) in an animal (e.g. gene
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 291 AA;
XX
XX Query Match 29.9%; Score 458; DB 7; Length 291;
XX Best Local Similarity 34.9%; Pred. No. 5.8e-39;
XX Matches 96; Conservative 57; Mismatches 104; Indels 18; Gaps 4;
XX
XX 24 GVRLPPFCRGKMEIANPPVRPSDVWIVTPYKSGTSLQEVYLVVSQGDPEIGLMNID 83
XX | : : : : : | : : : : : | : : : : : | : : : : : |
XX 13 GIPLIKYPAETIGPLQNFTAWPDDLLISTYPKSGTTWMSEILDIMYQGGKLEKCGRPIY 72
XX | : : : : : | : : : : : | : : : : : | : : : : : |
XX 84 EQLPVLEYPOP-----GLDIIKELTSPRLIKSHLPRFLPSDLHNGDSKVIYMARNPKDLV 139
XX | : : : : : | : : : : : | : : : : : | : : : : : |
XX 73 ARVPLEPKCGVSGSLTETBETPAPLKTHLPUSLPQSLDQKQKVIYIARNKDDV 132
XX | : : : : : | : : : : : | : : : : : | : : : : : |
XX 140 VSYQFHRSLRTMSYRGTFQFCFRFMNDKLGYSWFHFHVDSEFVFLFKYEDM 199
XX | : : : : : | : : : : : | : : : : : | : : : : : |
XX 133 VSYNYFNWAKLHPDPGTFWDSFLENFMDSVSGSYQHVKEWELRTHPVLVLFYEDI 192
XX | : : : : : | : : : : : | : : : : : | : : : : : |
XX 200 HRDLVTWVEQLARFLGVSCKAQLEALTEHC--HOLVDQC-CNAEALP-----V 245
XX | : : : : : | : : : : : | : : : : : | : : : : : |
XX 193 KENPKREIKKILEFLGRSLPSETVDSIVHSTSFKKMKENCMTNYTITPTIMDHNVSFPM 252
XX | : : : : : | : : : : : | : : : : : | : : : : : |
XX 246 GRGRVGLWKDIFTVSMNKKFDLVYKQKWKCDLTF 260
XX | : : : : : | : : : : : | : : : : : | : : : : : |
XX 253 RKGTTGDKWKTFTVAQNERFRAHYAKTWTDCDFK 287
XX
XX RESULT 22
XX ADE57135
XX ID ADE57135 standard; protein; 291 AA.
XX
```

```
AC ADE57135;
XX
XX 29-JAN-2004 (first entry)
XX
XX Rat Protein P17988, SEQ ID NO 2995.
XX
XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
XX chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
XX Rattus norvegicus.
XX
XX WO2003016475-A2.
XX
XX 27-FEB-2003.
XX
XX 14-AUG-2002; 2002WO-US025765.
XX
XX 14-AUG-2001; 2001US-0312147P.
XX 01-NOV-2001; 2001US-0346382P.
XX 26-NOV-2001; 2001US-033347P.
XX
XX (GHO ) GEN HOSPITAL CORP.
XX (FARB ) BAYER AG.
XX
XX Woolf C, D'urso D, Befort K, Costigan M;
XX
XX WPI; 2003-268312/26.
XX GENBANK; P17988.
XX
XX New composition comprising two or more isolated polypeptides, useful for
XX preparing a medicament for treating pain in an animal.
XX
XX Claim 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
XX or human polynucleotides or a polynucleotide which represents a fragment,
XX derivative or allelic variation of the nucleic acid sequence. Also
XX claimed are a vector comprising the novel polynucleotide, a host cell
XX comprising the vector, a method for identifying a nucleotide sequence
XX which is differentially regulated in an animal subjected to pain and a
XX kit to perform the method, an array, a method for identifying an agent
XX that increases or decreases the expression of the polynucleotide sequence
XX that is differentially expressed in neuronal tissue of a first animal
XX subjected to pain, a method for identifying a compound which regularly
XX the expression of a polynucleotide sequence which is differentially
XX expressed in an animal subjected to pain, a method for identifying a
XX compound that regulates the activity of one or more of the
XX polynucleotides, a method for producing a pharmaceutical composition, a
XX method for identifying a compound or small molecule that regulates the
XX activity in an animal of one or more of the polypeptides given in the
XX specification, a method for identifying a compound useful in treating
XX activity in an animal of one or more of the polypeptides given in the
XX specification, a method for producing a compound useful in treating
XX pain (e.g. spinal segmental nerve injury (SNI)) in an animal (e.g. gene
XX injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
XX therapy). The sequence presented is a rat protein (shown in Table 2 of
XX the specification) which is differentially expressed during pain. Note:
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic form directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 291 AA;
XX
XX Query Match 29.9%; Score 458; DB 7; Length 291;
XX Best Local Similarity 34.9%; Pred. No. 5.8e-39;
XX Matches 96; Conservative 57; Mismatches 104; Indels 18; Gaps 4;
XX
XX 24 GVRLPPFCRGKMEIANPPVRPSDVWIVTPYKSGTSLQEVYLVVSQGDPEIGLMNID 83
XX | : : : : : | : : : : : | : : : : : | : : : : : |
XX 13 GIPLIKYPAETIGPLQNFTAWPDDLLISTYPKSGTTWMSEILDIMYQGGKLEKCGRPIY 72
XX | : : : : : | : : : : : | : : : : : | : : : : : |
XX 84 EQLPVLEYPOP-----GLDIIKELTSPRLIKSHLPRFLPSDLHNGDSKVIYMARNPKDLV 139
XX | : : : : : | : : : : : | : : : : : | : : : : : |
XX 73 ARVPLEPKCGVSGSLTETBETPAPLKTHLPUSLPQSLDQKQKVIYIARNKDDV 132
XX | : : : : : | : : : : : | : : : : : | : : : : : |
XX 140 VSYQFHRSLRTMSYRGTFQFCFRFMNDKLGYSWFHFHVDSEFVFLFKYEDM 199
XX | : : : : : | : : : : : | : : : : : | : : : : : |
XX 133 VSYNYFNWAKLHPDPGTFWDSFLENFMDSVSGSYQHVKEWELRTHPVLVLFYEDI 192
XX | : : : : : | : : : : : | : : : : : | : : : : : |
XX 200 HRDLVTWVEQLARFLGVSCKAQLEALTEHC--HOLVDQC-CNAEALP-----V 245
XX | : : : : : | : : : : : | : : : : : | : : : : : |
XX 193 KENPKREIKKILEFLGRSLPSETVDSIVHSTSFKKMKENCMTNYTITPTIMDHNVSFPM 252
XX | : : : : : | : : : : : | : : : : : | : : : : : |
XX 246 GRGRVGLWKDIFTVSMNKKFDLVYKQKWKCDLTF 260
XX | : : : : : | : : : : : | : : : : : | : : : : : |
XX 253 RKGTTGDKWKTFTVAQNERFRAHYAKTWTDCDFK 287
XX
XX RESULT 22
XX ADE57135
XX ID ADE57135 standard; protein; 291 AA.
XX
```


XX The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 291 AA;

Query Match 29.9%; Score 458; DB 7; Length 291;
 Best Local Similarity 34.9%; Pred. No. 5.8e-39;
 Matches 96; Conservative 57; Mismatches 104; Indels 18; Gaps 4;

QY 24 GVRLPFCRGKMEETANFVRPSDVWIVTPKSGTSLQEVVYLVSQGADPDEIGLMNID 83
 DB 13 GIPLIKYFAETIGPQNFTAMPDDLLISTYPKSGTWTWSEILDMTYQGGKLEKCGRAPIY 72
 QY 84 EQLPVEYQPQ---GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNP KDVL 139
 DB 73 ARVPLEFKCQVPSGLETETPAPRLKTHLP LSLPQSLDQKVKVVIARNAKDV 132
 QY 140 VSYOYFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEBHQRMDSNVLFUKYEDM 199
 DB 133 VSYNYFNNAKLHPDPGTWDSPLENFMDSGVSYSWYQHVKEWELRTHPVLVLYFYEDI 192
 QY 200 HRDLVTWVEQLARFLGVSCDKAQLALTEHC--HQLVDQC-CNAALP-----V 245
 DB 193 KENPKREIKKLEFLGRSLPEETVDSIVHHTSFKKMKNCMTNTYTIPTIEMDHNVSPPM 252
 QY 246 GRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
 DB 253 RKGTTGDWNTFTVAQNERFDAHYAKTWD CDFK 287

RESULT 25

AAV67294

ID AAV67294 standard; protein; 295 AA.

XX AAV67294;

XX 05-APR-2000 (first entry)

XX Human STP2 (phenol sulphotransferase 2) amino acid sequence.

XX Single nucleotide polymorphism; SNP; STP2; phenol sulphotransferase;

XX probe; genotyping; human; drug metabolism.

XX Homo sapiens.

XX WO9964630-A1.

XX

PD 16-DEC-1999.
 XX
 PF 09-JUN-1999; 99WO-US013094.
 XX
 PR 10-JUN-1998; 98US-0088710P.
 XX
 PA (AXYS-) AXYS PHARM INC.
 XX
 PI Guida M, Kurth J;
 XX
 DR WPI; 2000-105892/09.
 XX
 DR N-PSDB; AAZ53953.
 XX
 PT Novel nucleic acid used for genotyping, e.g. to predict rate of drug
 PT metabolism.
 PS Disclosure; Page 28; 46pp; English.
 XX
 CC This is the human phenol sulphotransferase 2 (STP2) protein sequence. The
 CC invention relates to sequences AAZ5305-Z59352 which are fragments of the
 CC STP2 gene. The fragments are from the 8 exons, the promoter region, 3',
 CC and 5' untranslated regions of the STP2 gene. Each of the sequences
 CC contains a newly identified STP2 gene single nucleotide polymorphism
 CC (SNP). STP2 is a phenol sulphotransferase. Substrates for STP2 include
 CC minoxidil, acetaminophen, and paracetamol. Several of the nucleotide
 CC changes identified at the polymorphism sites, give rise to an amino acid
 CC change. Amino acid changes may result in altered enzyme activity. The
 CC sequences can be used as probes for detecting STP2 polymorphisms. The
 CC polymorphic probes are used in screening and genotyping, i.e. to predict
 CC the rate of metabolism of STP2 substrates, potential drug-drug
 CC interactions and adverse side effects. They can also be used to detect
 CC diseases resulting from accidental or occupational exposure to toxins and
 CC to establish animal, cell or in vitro models for drug metabolism
 XX
 SQ Sequence 295 AA;

Query Match 29.8%; Score 456; DB 3; Length 295;
 Best Local Similarity 33.7%; Pred. No. 9.6e-39;
 Matches 94; Conservative 58; Mismatches 101; Indels 26; Gaps 4;

QY 24 GVRLPFCRGKMEETANFVRPSDVWIVTPKSGTSLQEVVYLVSQGADPDEIGLMNID 83
 DB 17 GVPLIKYFAEALGPLQSFQARPDDLLISTYPKSGTWTWVQSLDMYQGGLEKCHRAPIF 76
 QY 84 EQLPVEYQPQ---GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNP KDVL 139
 DB 77 MRVPLEFKYVPGISGMETLKNTPAPRLKTHLP LALLPQTL DQKVKVVIARNAKDV 136
 QY 140 VSYOYFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEBHQRMDSNVLFUKYEDM 199
 DB 137 VSYNYFVHMAKVPHPGTWESFLEKFMAGEVSYGSWYQHVKEWELRTHPVLVLYFYEDM 196
 QY 200 HRDLVTWVEQLARFLGVSCDKAQLALTEHC-----HQLVDQCCNAEA 242
 DB 197 KENPKREIKKLEFLGRSLPEETVDSIVHHTSFKKMKNCMTNTYTIPTIEMDHNVSPPM 253
 QY 243 LP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
 DB 254 -PFWRKMGWAGDKTFTTVAQNERFDAHYAKTWD CDFK 291

RESULT 26
 ADP65305
 ID ADP65305 standard; protein; 295 AA.
 XX
 AC ADP65305;
 XX
 DT 12-AUG-2004 (first entry)
 XX
 DE Human sulphotransferase family, cytosolic, 1a.
 XX
 KW autoimmune disease; arthritis; collagen-induced; immunosuppressive; antirheumatic;
 KW rheumatoid arthritis; collagen-induced; immunosuppressive; antirheumatic;

KW antiarthritic; osteopathic; antigout; antiinflammatory; dermatological;
KW immunomodulatory; lupus; ankylosing spondylitis; Fibrositis;
KW fibromyalgia; osteoarthritis; gout; juvenile rheumatoid arthritis;
KW immune; human.
XX Homo sapiens.
XX WO2003072827-A1.
XX PD 04-SEP-2003.
XX PF 31-OCT-2002; 2002WO-US035433.
XX PR 31-OCT-2001; 2001US-0336220P.
XX PA (CHIL-) CHILDREN'S HOSPITAL MEDICAL CENT.
XX PI Hirsch R, Thorton SL;
XX DR WPI; 2003-712740/67.
XX DR GENBANK; NP_001045.
XX PT Diagnosing and analyzing autoimmune disease using gene expression
XX PT profiles and microarray technology, useful for diagnosing and treating
XX PT rheumatoid arthritis, lupus, fibrositis, osteoarthritis, fibromyalgia and
XX PT gout.
XX PS Disclosure; Page; 56pp; English.
XX CC The invention relates to a novel method for diagnosing and analysing
XX CC autoimmune disease or arthritides. The method comprises obtaining a
XX CC patient sample containing mRNA, analysing gene expression using the mRNA
XX CC that results in a gene expression signature of the mRNA, and using that
XX CC gene expression signature to diagnose or analyse the autoimmune disease
XX CC or arthritides in the patient, where gene expression of at least 60% of
XX CC the genes correlates with that of the gene signature. The invention
XX CC further comprises a treatment of rheumatoid arthritis; identification of
XX CC genes for targeting in the treatment of rheumatoid arthritis in a mammal
XX CC other than a mouse; diagnosis of rheumatoid arthritis in a mammal; an
XX CC array or gene chip, specific for rheumatoid arthritis; diagnosis or
XX CC analyses of autoimmune disease or rheumatoid arthritis; screening the
XX CC efficacy of a candidate drug in vitro for the treatment of collagen-
XX CC induced arthritis; and reducing the symptoms associated with collagen-
XX CC induced arthritis. The compositions of the invention have the following
XX CC activities: immunosuppressive, antirheumatic, antiarthritic, osteopathic,
XX CC antigout, antiinflammatory, dermatological, and immunomodulatory. The
XX CC methods and compositions of the present invention are useful for
XX CC diagnosing and treating autoimmune disease or arthritides, such as
XX CC rheumatoid arthritis, lupus, ankylosing spondylitis, fibrositis,
XX CC fibromyalgia, osteoarthritis, gout, juvenile rheumatoid arthritis, and an
XX CC immune disease caused by an infectious agent. This sequence represents a
XX CC protein sequence relating to the genes used in the analysis and treatment
XX CC of autoimmune diseases or arthritides. Note: This sequence is not shown
XX CC in the specification. It has been supplied in an electronic format from
XX CC WIPO.
SQ Sequence 295 AA;
Query Match 29.8%; Score 456; DB 7; Length 295;
Best Local Similarity 33.7%; Pred. No. 9.6e-39;
Matches 94; Conservative 58; Mismatches 101; Indels 26; Gaps 4;
QY 24 GVRLPFFCRGMEETANFPVPSDVWVITYPKSGTSLLOEVYLVSGADPDEIGLMNID 83
DB 17 GVPLIKYFAEALGPQSQFARPDDLLISTYPSKGTWVSQILDMYIQQGDLKCHRAPIF 76
QY 84 EQLPVLLEYQPQ----GLDIIKELTSPRIKSHLPYRFLPSDLHNGDSXVIYMARPKDLV 139
DB 77 MRVPLEFVKFCIPSGMETLKNTPAPRLKLTHTLPLALLPQTLLDQKVYVVARNAKDVA 136
QY 140 VSYQFHSRLTMSVYRGTFQFCRRFMDNKLGYSGWFHVSQFVHRHMDSNVLFKYEDM 199
DB 137 VSYTHFYHMAKYVPHPGTWESGFLEKFMAGEVSYGWYQHVQEWBLSRTHPVLVIFYEDM 196

QY 200 HRDLVTWVEQLARFLGVSCDKAQLALEATHEC-----HOLVDQCCNAEA 242
DB 197 KENPKRIQKILFEVGRSLPEETVDLMVHTSFEMKKKNPWTNTTVRRFMDHSIS--- 253
QY 243 LP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
DB 254 -PFMRKGMAGDWKTTFTVAQNERFDADYAEKMGACSLSF 291
RESULT 27
ADP24022
ID ADP24022 standard; protein; 295 AA.
XX AC ADP24022;
XX DT 18-NOV-2004 (first entry)
XX DE PRO polypeptide SEQ ID NO:1200.
XX KW PRO; antiinflammatory; antiarthritic; antirheumatic; immunosuppressive;
XX KW osteopathic; antidiabetic; dermatological; antipsoriatic; antiallergic;
XX KW antiaethmatic; hepatotropic; respiratory; gene therapy; immune system.
XX OS Unidentified.
XX PN WO2004041170-A2.
XX PD 21-MAY-2004.
XX PF 30-OCT-2003; 2003WO-US034312.
XX PR 01-NOV-2002; 2002US-0423394P.
XX PA (GETH) GENENTECH INC.
XX PI Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WT;
XX PI Wu TD;
XX DR WPI; 2004-419628/39.
XX DR N-PSDB; ADP24021.
XX PT New PRO polypeptides and polynucleotides, useful for treating e.g.
XX PT erythematosis, rheumatoid arthritis, diabetes mellitus, immune-mediated
XX PT renal disease, or demyelinating diseases of the central or peripheral
XX PT nervous system.
XX PS Claim 7; SEQ ID NO 1200; 2940pp; English.
XX CC The invention relates to a novel isolated nucleic acid and the PRO
XX CC polypeptide encoded by it. A protein of the invention has
XX CC antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,
XX CC osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,
XX CC antiaethmatic, hepatotropic, and respiratory activity. A polynucleotide
XX CC of the invention may have a use in gene therapy. The PRO polypeptide, its
XX CC agonist, antagonist, or antibody that specifically binds to the
XX CC polypeptide is useful for treating an immune related disorder such as
XX CC systemic lupus erythematosis, rheumatoid arthritis, osteoarthritis,
XX CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an
XX CC idiopathic inflammatory myopathy, Sjogren's syndrome, systemic
XX CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune
XX CC thrombocytopenia, chryoiditis, diabetes mellitus, immune-mediated renal
XX CC disease, a demyelinating disease of the central or peripheral nervous
XX CC system, idiopathic demyelinating polynuropathy, Guillain-Barre syndrome,
XX CC a chronic inflammatory demyelinating polynuropathy, a hepatobiliary
XX CC disease, infectious or autoimmune chronic active hepatitis, primary
XX CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,
XX CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's
XX CC disease, an autoimmune or immune-mediated skin disease, a bullous skin
XX CC disease, erythema multiforme, contact dermatitis, psoriasis, an allergic
XX CC disease, asthma, allergic rhinitis, atopic dermatitis, food
XX CC hypersensitivity, urticaria, an immunologic disease of the lung,
XX CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity

CC pneumonitis, a transplantation associated disease, graft rejection or
CC graft-versus-host disease. The present sequence represents a PRO protein
CC of the invention.
XX
XX
SQ Sequence 295 AA;
Query Match 29.8%; Score 456; DB 8; Length 295;
Best Local Similarity 33.7%; Pred. No. 9.6e-39;
Matches 94; Conservative 58; Mismatches 101; Indels 26; Gaps 4;
QY 24 GVRLPFCRGKMEETANFPVRPSDVWIVTPKSGTSLQEVVLYSQGADPDEIGLMNID 83
Db 17 GVPLIKYFAEALGPLOSFOARDDLLISTYPSKGTWVSQILDMYIQGGDLQKCHRAPIF 76
QY 84 EQLPVLEYPOP-----GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNP KDIV 139
Db 77 MRVPFLEPKVPGIPSGMETLKNTAPARLLKTHPLALLPQTLLDQKVYVYVARNAKDVA 136
QY 140 VSYIYOFHRSRLTMSYRGTFQFCRRFMNDKLGYSWFHVOEFWEHRMDSNVFLKYEDM 199
Db 137 VSYIYFHMAKYVPHPGTWESLEKFMAGEVSYGSWYQHVQEWELSRTHPVLYLYFYEDM 196
QY 200 HRDLVTMVEQLARFLGVSCDKAQLAALTEHC-----HQLVDQCCNAEA 242
Db 197 KENPKREIQKILEFVGRSLPEETVDLMVEHTSFKEKKNQMTNVTYVRRFMDHSIS--- 253
QY 243 LP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 254 -PFMRKGWAGDKTFTVAQNERFDADYAEKMGAGCSLSF 291

RESULT 28
ADW71800
ID ADW71800 standard; protein; 295 AA.
XX
XX AC ADW71800;
XX
XX DT 07-APR-2005 (first entry)
XX
XX DE Pig phenol sulfotransferase protein, SULT1A1, SEQ ID 1.
XX
XX KW Selectable marker; screening; animal breeding; polymorphism;
KW restriction fragment length polymorphism; allelic variation;
KW aryl sulfotransferase; PST; phenol sulfotransferase; SULT1A1; enzyme.
XX
XX OS Sus scrofa.
XX
XX FH Key Location/Qualifiers
FT Active-site 83
FT Active-site 134
FT Active-site 263
XX
XX PN US2005019788-A1.
XX
XX PD 27-JAN-2005.
XX
XX PF 30-JAN-2004; 2004US-00769507.
XX
XX PR 08-APR-1998; 98US-0081037P.
XX PR 08-APR-1999; 99US-00288037.
XX PR 23-NOV-2001; 2001US-00024628.
XX
XX PA (UYGU-) UNIV GUELPH.
XX
XX PI Squires EJ, Lin Z, Lou Y;
XX
XX DR N-PSDB; ADW71804.
XX
XX PT Genetically typing animals to determine those with desired boar taint
XX PT characteristics, comprises obtaining a sample of genetic material from
XX PT the animal, and assaying for the presence of a sulfotransferase allele.

PS Example; SEQ ID NO 1; 24pp; English.
XX
XX CC The present invention relates to a method of genetically typing animals
CC to determine those with desired boar taint characteristics. The method
CC involves obtaining a sample of genetic material from the animal and
CC assaying for the presence of a sulfotransferase allele. The present
CC sequence is the pig phenol sulfotransferase (PST) protein, SULT1A1.
XX
SQ Sequence 295 AA;
Query Match 29.8%; Score 456; DB 9; Length 295;
Best Local Similarity 35.6%; Pred. No. 9.6e-39;
Matches 98; Conservative 51; Mismatches 108; Indels 18; Gaps 3;
QY 24 GVRLPFCRGKMEETANFPVRPSDVWIVTPKSGTSLQEVVLYSQGADPDEIGLMNID 83
Db 17 GVPLIKYFAEALGPLESQAMPDDVLIISTYPSKGTWVSEILDLIYQGGDLQKCHRAPIF 76
QY 84 EQLPVLEYPOP-----GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNP KDIV 139
Db 77 MRVPFLEPKIPRCPTGFEKLDTPAPRLLKTHPLTLLPQTLLDQKVYVYVARNAKDVA 136
QY 140 VSYIYOFHRSRLTMSYRGTFQFCRRFMNDKLGYSWFHVOEFWEHRMDSNVFLKYEDM 199
Db 137 VSYIYFHMAKYVNPNGTWDSDLEDFMAGEVSYGSWYQHVQEWELSRTHPVLYLYFYEDM 196
QY 200 HRDLVTMVEQLARFLGVSCDKAQLAALTEHC---HQLVDQCCNAEALP-----V 245
Db 197 KENPKREIQKILEFVGRSLPEETVEDIVQHTSFQEMKNAMNTYTLPSDLLDHSISAFM 256
QY 246 GRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 257 RKGITGDKWKTFTVAQNERFEADYAEKMGAGCNLRF 291
RESULT 29
ADW71802
ID ADW71802 standard; protein; 295 AA.
XX
XX AC ADW71802;
XX
XX DT 07-APR-2005 (first entry)
XX
XX DE Human phenol sulfotransferase protein, SULT1A2, SEQ ID 3.
XX
XX KW Selectable marker; screening; animal breeding; polymorphism;
KW restriction fragment length polymorphism; allelic variation;
KW aryl sulfotransferase; PST; phenol sulfotransferase; SULT1A2; enzyme.
XX
XX OS Homo sapiens.
XX
XX FH Key Location/Qualifiers
FT Active-site 83
FT Active-site 134
FT Active-site 263
XX
XX PN US2005019788-A1.
XX
XX PD 27-JAN-2005.
XX
XX PF 30-JAN-2004; 2004US-00769507.
XX
XX PR 08-APR-1998; 98US-0081037P.
XX PR 08-APR-1999; 99US-00288037.
XX PR 23-NOV-2001; 2001US-00024628.
XX
XX PA (UYGU-) UNIV GUELPH.
XX
XX PI Squires EJ, Lin Z, Lou Y;
XX
XX DR WPI; 2005-111909/12.
XX
XX PT Genetically typing animals to determine those with desired boar taint

PT characteristics, comprises obtaining a sample of genetic material from
PT the animal, and assaying for the presence of a sulfotransferase allele.
XX
XX
PS Example; SEQ ID NO 3; 24pp; English.
XX
CC The present invention relates to a method of genetically typing animals
CC to determine those with desired boar taint characteristics. The method
CC involves obtaining a sample of genetic material from the animal and
CC assaying for the presence of a sulfotransferase allele. The present
CC sequence is the human phenol sulfotransferase (PST) protein, SULT1A2. The
CC SULT1A2 protein encoding gene is located on chromosome 16p12.1.
XX
SQ Sequence 295 AA;

Query Match 29.8%; Score 456; DB 9; Length 295;
Best Local Similarity 33.7%; Pred. No. 9.6e-39;
Matches 94; Conservative 58; Mismatches 101; Indels 26; Gaps 4;
QY 24 GVRLLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSLQEVVYLVSQGADPDEIGLMNID 83
DB 17 GVPLIKYFAEALGPQSQAPRDDLLISTYPKSGTTWVSQILDMMIYQGGDLKCHRAPIF 76
QY 84 EQLPVLEYPOP----GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNP KDLV 139
DB 77 MRVPLEFPKVPICPSGMETLKNTPAPRLKTHLPALLPQTLDDQKVKVYVARNAKDVA 136
QY 140 VSYIQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVOEFWEHRMDSNVLFKYEDM 199
DB 137 VSYHYFHYMAKVYPHPTGWESLEKFMAGEVSYSGSYQHVQEWSELSTHPTVLYLFYEDM 196
QY 200 HRDLVTWVEQLARFLGVSCDKAQLALTEHC-----HQLVDQCCNAEA 242
DB 197 KENPKREIQKILEFVGRSLPSETVDLMVEHTSFKEKKTPTWNTYTVRREFMDHSIS--- 253
QY 243 LP-VGRGVGLWKDIFTYSMNEKFDLVYKQKMGKCDLTF 280
DB 254 -PFMRKGMAGDWKTTFTVAQNERFDADYAEKMGAGCSLSF 291

RESULT 30
ADMW71805
ID ADMW71805 standard; protein; 295 AA.
XX ADMW71805;
XX
DT 07-APR-2005 (first entry)
XX
DE Pig phenol sulfotransferase protein, SULT1A1, SEQ ID 6.
XX
KW Selectable marker; screening; animal breeding; polymorphism;
KW restriction fragment length polymorphism; allelic variation;
KW aryl sulfotransferase; PST; phenol sulfotransferase; SULT1A1; enzyme.
XX
OS Sus scrofa.
XX
FH Key Location/Qualifiers
FT Active-site 83
FT Active-site 134
FT Active-site 263
XX
XX US2005019788-A1.
XX
PD 27-JAN-2005.
XX
PF 30-JAN-2004; 2004US-00769507.
XX
PR 08-APR-1998; 98US-0081037P.
PR 08-APR-1999; 99US-00288037.
PR 23-NOV-2001; 2001US-00024628.
XX
XX (UYGU-) UNIV GUELPH.
XX
XX Squires EJ, Lin Z, Lou Y;

XX WPI; 2005-111909/12.
DR N-FSDB; ADMW71804.
XX
PT Genetically typing animals to determine those with desired boar taint
PT characteristics, comprises obtaining a sample of genetic material from
PT the animal, and assaying for the presence of a sulfotransferase allele.
XX
PS Claim 18; SEQ ID NO 6; 24pp; English.
XX
CC The present invention relates to a method of genetically typing animals
CC to determine those with desired boar taint characteristics. The method
CC involves obtaining a sample of genetic material from the animal and
CC assaying for the presence of a sulfotransferase allele. The present
CC sequence is the pig phenol sulfotransferase (PST) protein, SULT1A1.
XX
SQ Sequence 295 AA;

Query Match 29.8%; Score 456; DB 9; Length 295;
Best Local Similarity 35.6%; Pred. No. 9.6e-39;
Matches 98; Conservative 51; Mismatches 108; Indels 18; Gaps 3;
QY 24 GVRLLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSLQEVVYLVSQGADPDEIGLMNID 83
DB 17 GVPLIKYFAEALGPLESFQAWPDVLISTYPKSGTTWVSEILDLYIQGGDLQKCORAPIF 76
QY 84 EQLPVLEYPOP----GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNP KDLV 139
DB 77 VRVPLEFPKIPROCTGPFELKQDTPAPRLKTHLPQTLDDQKVKVYVARNAKDVA 136
QY 140 VSYIQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVOEFWEHRMDSNVLFKYEDM 199
DB 137 VSYHYFHYMAKVYPHPTGWESLEKFMAGEVSYSGSYQHVQEWSELSTHPTVLYLFYEDM 196
QY 200 HRDLVTWVEQLARFLGVSCDKAQLALTEHC-----HQLVDQCCNAEA 245
DB 197 KENPKREIQKILEFVGRSLPSETVEDIVQHTSFQEMKNAMNTYRTLPDLDLHDSISAFM 256
QY 246 GRRGVGLWKDIFTYSMNEKFDLVYKQKMGKCDLTF 280
DB 257 RKGITGDMKSTFTVAQNERFEADYAEKMGAGCNLRF 291

RESULT 31
ADMW71815
ID ADMW71815 standard; protein; 295 AA.
XX ADMW71815;
XX
DT 07-APR-2005 (first entry)
XX
DE Pig phenol sulfotransferase mutant protein, SULT1A1.
XX
KW Selectable marker; screening; animal breeding; polymorphism;
KW restriction fragment length polymorphism; allelic variation;
KW aryl sulfotransferase; PST; phenol sulfotransferase; SULT1A1; enzyme;
KW mutein.
XX
OS Sus scrofa.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 147
FT /note= "Wild-type Lys substituted with Glu"
XX
XX US2005019788-A1.
XX
PD 27-JAN-2005.
XX
PF 30-JAN-2004; 2004US-00769507.
XX
PR 08-APR-1998; 98US-0081037P.
PR 08-APR-1999; 99US-00288037.

PR 23-NOV-2001; 2001US-00024628.
XX (UYGU-) UNIV GUELPH.
PA Squires EJ, Lin Z, Lou Y;
PI WPI: 2005-1111909/12.
DR N-PSDB; ADW71814.
XX Genetically typing animals to determine those with desired boar taint characteristics, comprises obtaining a sample of genetic material from the animal, and assaying for the presence of a sulfotransferase allele.
XX Example; Page; 24pp; English.
XX The present invention relates to a method of genetically typing animals to determine those with desired boar taint characteristics. The method involves obtaining a sample of genetic material from the animal and assaying for the presence of a sulfotransferase allele. The present sequence is the pig phenol sulfotransferase (PST) mutant protein, CC SULF1A1. Note: This sequence is not shown in specification but is derived from the pig wild-type phenol sulfotransferase shown as SEQ ID NO:6 in the specification.
XX Sequence 295 AA;
SQ
Query Match 29.7%; Score 454; DB 9; Length 295;
Best Local Similarity 35.6%; Pred. No. 1.6e-38;
Matches 98; Conservative 50; Mismatches 109; Indels 18; Gaps 3;
QY 24 GVRLPFFCRGKMEETANFVRPSDVIVTYPKSGTSLLOEVVYVLSQADPDEIGLMNID 83
DB 17 GVPLIKYFAEALGPLQSFQAPDDVLIISTYPKSGTWTWSEILDLYQGGDLQKQCAPIF 76
QY 84 EQLPVLEYPOP---GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYVARNPKDLV 139
DB 77 VRVPFLEFKIPRCPTGFGELLKDTAPAPRLKTHLPDLTPQTLDDOKVYVYVARNAKDVA 136
QY 140 VSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVFQFWEHRMDSNVFLFLKYEDM 199
DB 137 VSYHFYHMAEVYVNPFGTWSFLEDFMAGEVSGYSQYHVSQWELRHTHPVLYLFYEDM 196
QY 200 HRDLVTWVEQLARFLGVSCDKAQLAETHC---HQLVDOCCNAEALP-----V 245
DB 197 KENPKREIQKILEFVGRSLPEETVEDIVQHTSFQEMKKNAMTYRTLPDLLDHSISAFM 256
QY 246 GRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
DB 257 RKGITGDMKSTFTVAQNERFEADYAEKMGACNLRF 291
RESULT 32
ADD18776
ID ADD18776 standard; protein; 295 AA.
XX ADD18776;
XX 15-JAN-2004 (first entry)
XX Human disease related protein SeqID207.
XX human; disease state; cytostatic; antiinflammatory; ophthalmological;
KW antiarteriosclerotic; vulnary; gene therapy;
KW hypoxia-regulated condition; tumorigenesis; angiogenesis; apoptosis;
KW inflammation; erythropoiesis; glycolysis; gluconeogenesis;
KW glucose transportation; catecholamine synthesis; iron transport;
KW nitric oxide synthesis; cancer; ischaemic condition; reperfusion injury;
KW retinopathy; neonatal stress; pre-eclampsia; atherosclerosis;
KW inflammatory condition; wound healing.
XX Homo sapiens.
XX WO2003018621-A2.
PN

XX 06-MAR-2003.
XX 23-AUG-2002; 2002WO-GB003892.
XX 23-AUG-2001; 2001GB-00020558.
PR 05-OCT-2001; 2001GB-00024037.
XX (OXFO-) OXFORD BIOMEDICA UK LTD.
XX Kingsman SM, White J, Ward NR, Harris RA, Naylor S, Mundy CR;
PI WPI: 2003-290046/28.
DR N-PSDB; ADD18777.
XX New substantially purified polypeptide, useful for diagnosing or treating a hypoxia-regulated condition, such as cancer, ischemia, reperfusion injury, retinopathy, pre-eclampsia, atherosclerosis, inflammation, or wound healing.
XX Claim 25; SEQ ID NO 207; 424pp; English.
XX This invention relates to novel human genes and gene product which are implicated in certain disease states. Compounds which modulate the proteins of the invention may have cytostatic, antiinflammatory, ophthalmological, antiarteriosclerotic or vulnary activities. The sequences of the invention may be useful for gene therapy. The invention may be useful for diagnosing or treating a hypoxia-regulated condition, such as tumorigenesis, angiogenesis, apoptosis, inflammation, erythropoiesis, or the biological response to hypoxia conditions including processes such as glycolysis, gluconeogenesis, glucose transportation, catecholamine synthesis, iron transport or nitric oxide synthesis. The disease includes cancer, ischaemic conditions, reperfusion injury, retinopathy, neonatal stress, pre-eclampsia, atherosclerosis, inflammatory conditions or wound healing. The present sequence is that of a disease related protein of the invention.
XX Sequence 295 AA;
SQ
Query Match 29.5%; Score 451; DB 7; Length 295;
Best Local Similarity 33.8%; Pred. No. 3.2e-38;
Matches 93; Conservative 57; Mismatches 107; Indels 18; Gaps 3;
QY 24 GVRLPFFCRGKMEETANFVRPSDVIVTYPKSGTSLLOEVVYVLSQADPDEIGLMNID 83
DB 17 GVPLIKYFAEALGPLQSFQAPDDVLIISTYPKSGTWTWSEILDLYQGGDLQKCHRAPIF 76
QY 84 EQLPVLEYPOP---GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYVARNPKDLV 139
DB 77 MRVPFLEFKAPGIPSGMETLKTDPAPRLKTHLPDLTPQTLDDOKVYVYVARNAKDVA 136
QY 140 VSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVFQFWEHRMDSNVFLFLKYEDM 199
DB 137 VSYHFYHMAEVYVNPFGTWSFLEKFVNGEVSQYHVSQWELRHTHPVLYLFYEDM 196
QY 200 HRDLVTWVEQLARFLGVSCDKAQLAETHC---HQLVDOCCNAEALP-----V 245
DB 197 KENPKREIQKILEFVGRSLPEETVEDIVQHTSFQEMKKNAMTYRTVTPQEFMDHSISPFM 256
QY 246 GRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
DB 257 RKGWAGDMKSTFTVAQNERFDADYAEKMGACLSLF 291
RESULT 33
ADE57149
ID ADE57149 standard; protein; 295 AA.
XX ADE57149;
XX 29-JAN-2004 (first entry)
XX Human Protein P50225, SEQ ID NO 3009.
DE

```
XX Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
XX
OS Homo sapiens.
XX
XX WO2003016475-A2.
XX
XX PD 27-FEB-2003.
XX
XX PF 14-AUG-2002; 2002WO-US025765.
XX
XX PR 14-AUG-2001; 2001US-0312147P.
XX
XX PR 01-NOV-2001; 2001US-0346382P.
XX
XX PR 26-NOV-2001; 2001US-0333347P.
XX
XX PA (GEHO ) GEN HOSPITAL CORP.
XX
XX PA (FARB ) BAYER AG.
XX
XX PI Woolf C, D'urso D, Befort K, Costigan M;
XX
XX DR WPI; 2003-268312/26.
XX
XX DR GENBANK; P50225.
XX
XX PT New composition comprising two or more isolated polypeptides, useful for
XX preparing a medicament for treating pain in an animal.
XX
XX PS Claim 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
XX or human polynucleotides or a polynucleotide which represents a fragment,
XX derivative or allelic variation of the nucleic acid sequence. Also
XX claimed are a vector comprising the novel polynucleotide, a host cell
XX comprising the vector, a method for identifying a nucleotide sequence
XX which is differentially regulated in an animal subjected to pain and a
XX kit to perform the method, an array, a method for identifying an agent
XX that increases or decreases the expression of the polynucleotide sequence
XX that is differentially expressed in neuronal tissue of a first animal
XX subjected to pain, a method for identifying a compound which regulates
XX the expression of a polynucleotide sequence which is differentially
XX expressed in an animal subjected to pain, a method for identifying a
XX compound that regulates the activity of one or more of the
XX polynucleotides, a method for producing a pharmaceutical composition, a
XX method for identifying a compound or small molecule that regulates the
XX activity in an animal of one or more of the polypeptides given in the
XX specification, a method for identifying a compound useful in treating
XX pain and a pharmaceutical composition comprising the one or more
XX polypeptides or their antibodies. The polynucleotide or the compound that
XX modulates its activity is useful for preparing a medicament for treating
XX pain (e.g. spinal segmental nerve injury (SNI)) in an animal (e.g. gene
XX therapy). The sequence presented is a human protein (shown in Table 2 of
XX the specification) which is differentially expressed during pain. Note:
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic form directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 295 AA;
XX
XX Query Match
XX Best Local Similarity 29.5%; Score 451; DB 7; Length 295;
XX Matches 93; Conservative 57; Mismatches 107; Indels 18; Gaps 3;
XX
XX QY 24 GVRLLPPFCRGKMEETANFVRPSDWIVITPKSGTSLIQEVYLVVSQGDPEIGLMNID 83
XX ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
XX 17 GVPLIKYFAEALGPLQSFQARDDLLISTPKSGTITWVSQILDMYQGDLEKCHRAPTF 76
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX QY 84 EQLPVLEYQPQ---GLDIIKELTSPRLTKSHLPYRFLPSDLHNGDSKVIYMARPKDLIV 139
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX DB 77 MKRVFLEKAPGIPSGMPTKPTAPRLKTHLPALPQTLGDQKVKVYVARNAKVA 136
XX
XX QY 140 VSYQFHRSLRTMSYRGTPQFCRRFMNDKLGYSWFHVEHWRMDSNVLFKLYEDM 199
XX
137 VSYHYFVHMAKVHPPEPGTWSDFLEKPMVGVSYSQVQHWELSKRTHPVLVLFYEDM 196
138 ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
139 HRDLVTWVEQLARPLGVSCDKAQLALTEHC---HQLVDQCCNAELP-----V 245
140 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
141 KENPKREIQKILEFVGRSLPEETVDFMVQHTSFKEMKKNPMTNTTVPQEFMDHSISPFM 256
142 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
143 GRGRVGLWKDIFTYSMNEKFDLVYKQKMGKCDLTF 280
144 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
145 RKGMDGDKTFTTVAQNERFDADYAEKMGAGCSLSF 291
146 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
RESULT 34
ADE57137
ID ADE57137 standard; protein; 295 AA.
XX
XX AC ADE57137;
XX
XX DT 29-JAN-2004 (first entry)
XX
XX DE Human Protein P50225, SEQ ID NO 2997.
XX
XX KW Human; pain; neuronal tissue; gene therapy;
XX KW spinal segmental nerve injury; chronic constriction injury; CCI;
XX KW spared nerve injury; SNI; Chung.
XX
XX OS Homo sapiens.
XX
XX PN WO2003016475-A2.
XX
XX PD 27-FEB-2003.
XX
XX PF 14-AUG-2002; 2002WO-US025765.
XX
XX PR 14-AUG-2001; 2001US-0312147P.
XX
XX PR 01-NOV-2001; 2001US-0346382P.
XX
XX PR 26-NOV-2001; 2001US-0333347P.
XX
XX PA (GEHO ) GEN HOSPITAL CORP.
XX
XX PA (FARB ) BAYER AG.
XX
XX PI Woolf C, D'urso D, Befort K, Costigan M;
XX
XX DR WPI; 2003-268312/26.
XX
XX DR GENBANK; P50225.
XX
XX PT New composition comprising two or more isolated polypeptides, useful for
XX preparing a medicament for treating pain in an animal.
XX
XX PS Claim 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
XX or human polynucleotides or a polynucleotide which represents a fragment,
XX derivative or allelic variation of the nucleic acid sequence. Also
XX claimed are a vector comprising the novel polynucleotide, a host cell
XX comprising the vector, a method for identifying a nucleotide sequence
XX which is differentially regulated in an animal subjected to pain and a
XX kit to perform the method, an array, a method for identifying an agent
XX that increases or decreases the expression of the polynucleotide sequence
XX that is differentially expressed in neuronal tissue of a first animal
XX subjected to pain, a method for identifying a compound which regulates
XX the expression of a polynucleotide sequence which is differentially
XX expressed in an animal subjected to pain, a method for identifying a
XX compound that regulates the activity of one or more of the
XX polynucleotides, a method for producing a pharmaceutical composition, a
XX method for identifying a compound or small molecule that regulates the
XX activity in an animal of one or more of the polypeptides given in the
XX specification, a method for identifying a compound useful in treating
XX pain and a pharmaceutical composition comprising the one or more
XX polypeptides or their antibodies. The polynucleotide or the compound that
XX modulates its activity is useful for preparing a medicament for treating
XX pain (e.g. spinal segmental nerve injury (SNI)) in an animal (e.g. gene
XX therapy). The sequence presented is a human protein (shown in Table 2 of
XX the specification) which is differentially expressed during pain. Note:
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic form directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 295 AA;
XX
XX Query Match
XX Best Local Similarity 29.5%; Score 451; DB 7; Length 295;
XX Matches 93; Conservative 57; Mismatches 107; Indels 18; Gaps 3;
XX
XX QY 24 GVRLLPPFCRGKMEETANFVRPSDWIVITPKSGTSLIQEVYLVVSQGDPEIGLMNID 83
XX ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
XX 17 GVPLIKYFAEALGPLQSFQARDDLLISTPKSGTITWVSQILDMYQGDLEKCHRAPTF 76
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX QY 84 EQLPVLEYQPQ---GLDIIKELTSPRLTKSHLPYRFLPSDLHNGDSKVIYMARPKDLIV 139
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX DB 77 MKRVFLEKAPGIPSGMPTKPTAPRLKTHLPALPQTLGDQKVKVYVARNAKVA 136
XX
XX QY 140 VSYQFHRSLRTMSYRGTPQFCRRFMNDKLGYSWFHVEHWRMDSNVLFKLYEDM 199
XX
```

CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 295 AA;

Query Match 29.5%; Score 451; DB 7; Length 295;
Best Local Similarity 33.8%; Pred. No. 3.2e-38;
Matches 93; Conservative 57; Mismatches 107; Indels 18; Gaps 3;

QY 24 GVRLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSLQEVVYLVSQGADPDEIGLMNID 83
DB 17 GVPLIKYFAEALGPLQSFQARPDDLLISTYPSKGTWVSQILDMIYQGGDLKCHRAPIF 76

QY 84 EQLPVLEYPOP---GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIVYMARNPOLV 139
DB 77 MRVPFLEFKAPGIPSGMETLKTDPAPRLKTHPLALLPQTLDDQKVYVYVARNAKVA 136

QY 140 VSYIYQFHRSLRTMSYRGTFQFCRFRPMNDKLGYSWFHVFQEFWEHRMDSNVLFKYEDM 199
DB 137 VSYHYFYHMAKVHPPEGTWDSFLEKFMVGEVSGSWYQHVQEWELSRTHPVLVLYFYEDM 196

QY 200 HRDLVTWVQLARFLGVSCDKAQLAETHC---HQLVQCCNAEALP-----V 245
DB 197 KENPKREIQKILEFVGRSLPSETVDFPMVQHTSFKEKMKNPMTNVTVPQEFMDHSISPPM 256

QY 246 GRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
DB 257 RKGMDGDKWTKTTTVAQNERFDADYAEKMGAGCSLSF 291

RESULT 35
ADE57141
ID ADE57141 standard; protein; 295 AA.
XX
AC ADE57141;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human Protein P50225, SEQ ID NO 3001.
XX
KW Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
OS Homo sapiens.
XX
PN WO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
XX (GHEO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
XX Woolf C, D'urso D, Befort K, Costigan M;
XX
XX WPI; 2003-268312/26.
DR GENBANK; P50225.
XX
XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
XX Claim 1; Page; 1017pp; English.
PS
XX The invention discloses a composition comprising two or more isolated rat

CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 295 AA;

Query Match 29.5%; Score 451; DB 7; Length 295;
Best Local Similarity 33.8%; Pred. No. 3.2e-38;
Matches 93; Conservative 57; Mismatches 107; Indels 18; Gaps 3;

QY 24 GVRLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSLQEVVYLVSQGADPDEIGLMNID 83
DB 17 GVPLIKYFAEALGPLQSFQARPDDLLISTYPSKGTWVSQILDMIYQGGDLKCHRAPIF 76

QY 84 EQLPVLEYPOP---GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIVYMARNPOLV 139
DB 77 MRVPFLEFKAPGIPSGMETLKTDPAPRLKTHPLALLPQTLDDQKVYVYVARNAKVA 136

QY 140 VSYIYQFHRSLRTMSYRGTFQFCRFRPMNDKLGYSWFHVFQEFWEHRMDSNVLFKYEDM 199
DB 137 VSYHYFYHMAKVHPPEGTWDSFLEKFMVGEVSGSWYQHVQEWELSRTHPVLVLYFYEDM 196

QY 200 HRDLVTWVQLARFLGVSCDKAQLAETHC---HQLVQCCNAEALP-----V 245
DB 197 KENPKREIQKILEFVGRSLPSETVDFPMVQHTSFKEKMKNPMTNVTVPQEFMDHSISPPM 256

QY 246 GRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
DB 257 RKGMDGDKWTKTTTVAQNERFDADYAEKMGAGCSLSF 291

RESULT 36
ADE57145
ID ADE57145 standard; protein; 295 AA.
XX
AC ADE57145;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human Protein P50225, SEQ ID NO 3005.
XX
KW Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
OS Homo sapiens.
XX
PN WO2003016475-A2.
XX
PD 27-FEB-2003.

XX PF 14-AUG-2002; 2002WO-US025765.
 XX PR 14-AUG-2001; 2001US-0312147P.
 PR 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-0333347P.
 XX PA (GBHO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.
 XX PI Woolf C, D'urso D, Befort K, Costigan M;
 XX WPI; 2003-268312/26.
 DR GENBANK; P50225.
 XX New composition comprising two or more isolated polypeptides, useful for
 PT preparing a medicament for treating pain in an animal.
 XX Claim 1; Page; 101pp; English.
 XX The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (SNRI)) in an animal (e.g. gene
 CC injury (CCI) and spared nerve injury (SNI)) in a human protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 295 AA;

Query Match 29.5%; Score 451; DB 7; Length 295;
 Best Local Similarity 33.8%; Pred. No. 3.2e-38;
 Matches 93; Conservative 57; Mismatches 107; Indels 18; Gaps 3;

QY 24 GVRLLPPFCRGKMEETANFVRPSDVWIVTPKSGTSLLOEVVYLVSQGADPDEIGLMNID 83
 DB 17 GVPLIKYFAEALGPLQSFQARPDLLISTYPKSGTTWVSQILDMTYQGDLKCHRAPIF 76
 QY 84 EQLPVLVEYQPQ---GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYVARNPKDLV 139
 DB 77 MRVPLEFKAQIPSGMETLKDTPTAPRLKTHLPALLPQTLLDQKVVVYVARNAKDVA 136
 QY 140 VSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVFQWFEHRMDSNVLFKLYEDM 199
 DB 137 VSYHYFHYHMAKVHPBPQWDSFLEKFMVGEVSGYQVQVSWELSRTHPVLYLFYEDM 196
 QY 200 HRDLVTWVEQLARFLGVSCDKAQLALETCHC---HQLVDQCCNAALP-----V 245
 DB 197 KENPKREIQKILEFVGRSLPFEETVDFMVQHTSFKEKMKKPNMTNTYTVQEFMDHSISPPM 256
 QY 246 GRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
 DB 257 RKGMAGDWKTFTTVAQNERFDADYAEKMGACSLSP 291

RESULT 37

ADI15901

XX ADI15901 standard; protein; 295 AA.

AC ADI15901;

XX DT 22-APR-2004 (first entry)

XX DE Human PP 88.

XX KW tumour; antigen; human leukocyte antigen; HLA-A2; HLA-A26;
 KW cytotoxic T-cell; CTL; vaccine; cancer; colorectal cancer;
 KW stomach cancer; buccal cancer; renal cancer; lung cancer;
 KW gynecological cancer; prostate cancer.

XX OS Homo sapiens.

XX PN WO2003008450-A1.

XX PD 30-JAN-2003.

XX PF 11-JUN-2002; 2002WO-JP005799.

XX PR 12-JUN-2001; 2001JP-00177058.

XX PR 21-AUG-2001; 2001JP-00250728.

XX PA (ITOH/) ITOH K.

XX PI Itoh K, Shichijo S;

XX DR WPI; 2003-267996/26.

XX DR N-PSDB; ADI15976.

XX Tumor antigen peptides recognized by human leukocyte antigen (HLA)-A2 or
 PT HLA-A26 restricted cytotoxic T-cells for treatment and prevention of
 PT cancer including preparation of cancer vaccines.

XX PS Claim 2; SEQ ID NO 245; 323pp; Japanese.

XX The invention relates to a tumour antigen peptide recognised by human
 CC leukocyte antigen (HLA)-A2 or HLA-A26 restricted cytotoxic T-cells (CTL)
 CC and/or capable of inducing CTL. The tumour antigen peptide is useful for
 CC the treatment, prevention, diagnosis and vaccine production for cancers
 CC including colorectal, stomach, buccal, renal, lung, gynecological and
 CC prostate cancer. The present sequence represents the amino acid sequence
 CC of a human protein.

XX SQ Sequence 295 AA;

Query Match 29.5%; Score 451; DB 7; Length 295;

Best Local Similarity 33.8%; Pred. No. 3.2e-38;

Matches 93; Conservative 57; Mismatches 107; Indels 18; Gaps 3;

QY 24 GVRLLPPFCRGKMEETANFVRPSDVWIVTPKSGTSLLOEVVYLVSQGADPDEIGLMNID 83

DB 17 GVPLIKYFAEALGPLQSFQARPDLLISTYPKSGTTWVSQILDMTYQGDLKCHRAPIF 76

QY 84 EQLPVLVEYQPQ---GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYVARNPKDLV 139

DB 77 MRVPLEFKAQIPSGMETLKDTPTAPRLKTHLPALLPQTLLDQKVVVYVARNAKDVA 136

QY 140 VSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVFQWFEHRMDSNVLFKLYEDM 199

DB 137 VSYHYFHYHMAKVHPBPQWDSFLEKFMVGEVSGYQVQVSWELSRTHPVLYLFYEDM 196

QY 200 HRDLVTWVEQLARFLGVSCDKAQLALETCHC---HQLVDQCCNAALP-----V 245

DB 197 KENPKREIQKILEFVGRSLPFEETVDFMVQHTSFKEKMKKPNMTNTYTVQEFMDHSISPPM 256

QY 246 GRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280

Db	257	RKGMAGDMKTTTVAQNERFDADYAEKMGCSLSF	291
RESULT 38			
ID	ADP23588		
XX	ADP23588	standard; protein; 295 AA.	
AC	ADP23588;		
DT	18-NOV-2004	(first entry)	
DE	PRO	polypeptide SEQ ID NO:766.	
KW	PRO;	antiinflammatory; antiarthritic; antirheumatic; immunosuppressive; osteopathic; antidiabetic; dermatological; antipsoriatic; antiallergic; antiasthmatic; hepatotropic; respiratory; gene therapy; immune system.	
XX	OS	Unidentified.	
XX	PN	WO2004041170-A2.	
XX	PD	21-MAY-2004.	
XX	PF	30-OCT-2003; 2003WO-US034312.	
XX	PR	01-NOV-2002; 2002US-0423394P.	
XX	PA	(GETH) GENENTECH INC.	
XX	PI	Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI; Wu TD;	
XX	XX	WPI; 2004-419628/39.	
DR	N-PSDB; ADP23587.		
XX	PT	New PRO polypeptides and polynucleotides, useful for treating e.g. erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated renal disease, or demyelinating diseases of the central or peripheral nervous system.	
XX	PS	Claim 7; SEQ ID NO 766; 2940pp; English.	
XX	CC	The invention relates to a novel isolated nucleic acid and the PRO polypeptide encoded by it. A protein of the invention has antiinflammatory, antiarthritic, antirheumatic, immunosuppressive, osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic, antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide of the invention may have a use in gene therapy. The PRO polypeptide, its agonist, antagonist, or antibody that specifically binds to the polypeptide is useful for treating an immune related disorder such as systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an idiopathic inflammatory myopathy, Sjogren's syndrome, systemic vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal disease, a demyelinating disease of the central or peripheral nervous system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome, a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary disease, infectious or autoimmune chronic active hepatitis, primary biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis, inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's disease, an autoimmune or immune-mediated skin disease, a bullous skin disease, erythema multiforme, contact dermatitis, psoriasis, an allergic disease, asthma, allergic rhinitis, atopic dermatitis, food hypersensitivity, urticaria, an immunologic disease of the lung, eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity pneumonitis, a transplantation associated disease, graft rejection or of the invention.	
XX	SQ	Sequence 295 AA;	
Query Match	29.5%;	Score 451; DB 8; Length 295;	

Best Local Similarity	33.8%;	Pred. No. 3.2e-38;	
Matches	93;	Conservative	57; Mismatches 107; Indels 18; Gaps 3;
QY	24	GVRLPFFCRGRMEETANFVRPDSVWIVTYPKSGTSLQEVYLVSQGADPDEIGLMNID	83
Db	17	GVPLIKYFAEALGPLQSFQARPDLLISTYPKSGTTWVSQILDMIYQGDLEKCHRAPIF	76
QY	84	EQLPVLVEYPOP-----GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYARNPKDLV	139
Db	77	MRVPFLEFAFGIPSGMETLKDTPAPRLTKTHLPALLPOTLLDQKVYVYVARNAKOVA	136
QY	140	VSYQFHRSLRTMSYRGTFQBFRCRFRFMDNLGYSWFEHVOEFWEHRMDSNVLFLLKYEDM	199
Db	137	VSYHYFTHMAKVHPEPGTWDSFLEKFMVGEVSYGSIYORVQEWELSRTHPVLILFYEDM	196
QY	200	HRDLVTWVEQLARFLGVSCDKAQLAEALTEHC---HQLVDQCCNAEALP-----V	245
Db	197	KENPKREIQKILEFVGHSLPBETVDFMVQHTSFKEKMKKNPMTNYTTVQEFMDHSISPM	256
QY	246	GRGRVGLWKDIFTVSMNKEKFDLVYKQXMGKCDLTF	280
Db	257	RKGMAGDMKTTTVAQNERFDADYAEKMGCSLSF	291
RESULT 39			
ID	ADR97362	standard; protein; 295 AA.	
XX	AC	ADR97362;	
XX	DT	02-DEC-2004	(first entry)
XX	DE	Human SULT1A1 protein, an apoptosis related target	Seq 70.
XX	KW	human; enzyme; apoptosis; cancer; inflammation; autoimmune; neurodegenerative disorder; cytostatic; antiinflammatory; immunosuppressive; neuroprotective; gene therapy; SULT1A1; sulfotransferase family.	
XX	OS	Homo sapiens.	
XX	PN	WO2004078783-A2.	
XX	PD	16-SEP-2004.	
XX	PF	05-MAR-2004; 2004WO-GB000957.	
XX	PR	07-MAR-2003; 2003GB-00005267.	
XX	PA	(EIRX-) EIRX THERAPEUTICS LTD.	
XX	PI	Murphy FJ, Sheehan DE, Keating KE, Hayes I, Seera L;	
XX	DR	WPI; 2004-662402/64.	
XX	DR	N-PSDB; ADR97361.	
XX	PT	Identifying an agent that modulates the function of an apoptosis-associated polypeptide, useful for diagnosing or treating e.g. cancer, comprises comparing the binding of the polypeptide to the candidate agent and to a control agent.	
XX	PS	Claim 1; SEQ ID NO 70; 304pp; English.	
XX	CC	This invention relates to novel agents that modulates the function of human apoptosis-associated proteins specified within the specification. Specifically, it refers to a method for the identification of target genes whose expression is correlated with an early stage in the regulation of apoptosis. The present invention describes a method of contacting either candidate agents or control agents to the target genes and assessing the difference of binding and inhibitory activity, where the candidate agent is selected from a low molecular weight organic molecule, an antibody or fragment thereof, an antisense oligonucleotide, a small inhibitory dsRNA, or a ribozyme. As such, the compositions and	

CC methods are useful for diagnosing and treating diseases or conditions
CC associated with abnormal apoptosis in mammalian tissue, such as cancer,
CC inflammation, autoimmune or neurodegenerative disorders. Accordingly,
CC they exhibit cytostatic, antiinflammatory, immunosuppressive and
CC neuroprotective activities. These may also be used for drug screening
CC purposes and in gene therapy. This polypeptide sequence is a human target
CC protein, an enzyme associated with the regulation of apoptosis whose
CC expression is modulated by novel agents of the invention.

XX Sequence 295 AA;

SQ Query Match 29.5%; Score 451; DB 8; Length 295;
Best Local Similarity 33.8%; Pred. No. 3.2e-38;
Matches 93; Conservative 57; Mismatches 107; Indels 18; Gaps 3;
QY 24 GVRLLPPFCRGKMEETANFVRPSDVWIVTPKSGTSLLOEVVYLVSQGADPDEIGLMNID 83
DB 17 GVPLIKYFAEALGPLQSFQARPDLLISTYPSKGTWVSQILDMIYQGGDLKCHRAPIF 76
QY 84 EQLPVLEYPOP---GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARPKDLV 139
DB 77 MRVPFLFKAPGIPSGMETLKTDPAPRLKTHLPALLPQTLLDQKVKVYVARNAKDVA 136
QY 140 VSYVQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHGFQFWEHRMDSNVFLFKYEDM 199
DB 137 VSYVHFYHMAKVHPPEPGTWSFLEKFMVGEVSGSWYQHVQEWELSRTHPVLYLYFYEDM 196
QY 200 HRDLVTMVEQLARFLGVSCDKAQLALTEHC---HQLVDQCCNAEALP-----V 245
DB 197 KENPKREIQKILEFVGHSLPEETVDFMVQHTSFKEKMKNPMTNTYTPQEFMDHSISPFM 256
QY 246 GRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
DB 257 RKGWAGDWKTTTVAQNERFDADYAEKMGAGCSLSF 291

RESULT 40

ADW71801
ID ADW71801 standard; protein; 295 AA.

XX AC ADW71801;

XX DT 07-APR-2005 (first entry)

XX DE Human phenol sulfotransferase protein, SULT1A1, SEQ ID 2.

XX KW Selectable marker; screening; animal breeding; polymorphism;
KW restriction fragment length polymorphism; allelic variation;
KW aryl sulfotransferase; PST; phenol sulfotransferase; SULT1A1; enzyme.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Active-site 83

FT Active-site 134

FT Active-site 263

XX US2005019788-A1.

XX 27-JAN-2005.

XX 30-JAN-2004; 2004US-00769507.

XX 08-APR-1998; 98US-0081037P.

XX 08-APR-1999; 99US-00288037.

XX 23-NOV-2001; 2001US-00024628.

XX (UYGU-) UNIV GUELPH.

XX Squires EJ, Lin Z, Lou Y;

XX WPI; 2005-111909/12.

XX

PT Genetically typing animals to determine those with desired boar taint
PT characteristics, comprises obtaining a sample of genetic material from
PT the animal, and assaying for the presence of a sulfotransferase allele.

XX Example; SEQ ID NO 2; 24pp; English.

XX The present invention relates to a method of genetically typing animals
CC to determine those with desired boar taint characteristics. The method
CC involves obtaining a sample of genetic material from the animal and
CC assaying for the presence of a sulfotransferase allele. The present
CC sequence is the human phenol sulfotransferase (PST) protein, SULT1A1. The
CC SULT1A1 protein encoding gene is located on chromosome 16p12.1.

SQ Sequence 295 AA;

Query Match 29.5%; Score 451; DB 9; Length 295;
Best Local Similarity 33.8%; Pred. No. 3.2e-38;
Matches 93; Conservative 57; Mismatches 107; Indels 18; Gaps 3;
QY 24 GVRLLPPFCRGKMEETANFVRPSDVWIVTPKSGTSLLOEVVYLVSQGADPDEIGLMNID 83
DB 17 GVPLIKYFAEALGPLQSFQARPDLLISTYPSKGTWVSQILDMIYQGGDLKCHRAPIF 76
QY 84 EQLPVLEYPOP---GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARPKDLV 139
DB 77 MRVPFLFKAPGIPSGMETLKTDPAPRLKTHLPALLPQTLLDQKVKVYVARNAKDVA 136
QY 140 VSYVQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHGFQFWEHRMDSNVFLFKYEDM 199
DB 137 VSYVHFYHMAKVHPPEPGTWSFLEKFMVGEVSGSWYQHVQEWELSRTHPVLYLYFYEDM 196
QY 200 HRDLVTMVEQLARFLGVSCDKAQLALTEHC---HQLVDQCCNAEALP-----V 245
DB 197 KENPKREIQKILEFVGHSLPEETVDFMVQHTSFKEKMKNPMTNTYTPQEFMDHSISPFM 256
QY 246 GRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
DB 257 RKGWAGDWKTTTVAQNERFDADYAEKMGAGCSLSF 291

RESULT 41

AD117050

ID AD117050 standard; protein; 302 AA.

XX AC AD117050;

XX DT 15-APR-2004 (first entry)

XX DE Human NOVX protein homologue SeqID 586.

XX KW human; NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes;
KW inflammation; autoimmune disorder; allergy; blood disorder;
KW acquired immunodeficiency syndrome; AIDS; obesity; asthma;
KW immunoglobulin (Ig)A nephropathy; cirrhosis; arthritis;
KW Alzheimer's disease; infection; str.

XX OS Homo sapiens.

XX WO200268649-A2.

XX 06-SEP-2002.

XX 31-JAN-2002; 2002WO-US0002785.

XX 31-JAN-2001; 2001US-0285395P.

XX 31-JAN-2001; 2001US-0265412P.

XX 31-JAN-2001; 2001US-0265514P.

XX 31-JAN-2001; 2001US-0265517P.

XX 02-FEB-2001; 2001US-0266406P.

XX 05-FEB-2001; 2001US-0266767P.

XX 07-FEB-2001; 2001US-0266975P.

XX 07-FEB-2001; 2001US-0267057P.

XX 08-FEB-2001; 2001US-0267459P.

09-FEB-2001; 2001US-0267823P.
15-FEB-2001; 2001US-0268974P.
26-FEB-2001; 2001US-0271664P.
27-FEB-2001; 2001US-0271839P.
27-FEB-2001; 2001US-0271855P.
02-MAR-2001; 2001US-0272788P.
02-MAR-2001; 2001US-0273046P.
14-MAR-2001; 2001US-0275252P.
14-MAR-2001; 2001US-0275947P.
14-MAR-2001; 2001US-0275950P.
15-MAR-2001; 2001US-0275989P.
15-MAR-2001; 2001US-0276448P.
15-MAR-2001; 2001US-0276450P.
16-MAR-2001; 2001US-0276397P.
16-MAR-2001; 2001US-0276768P.
20-MAR-2001; 2001US-0278652P.
26-MAR-2001; 2001US-0278775P.
26-MAR-2001; 2001US-0278778P.
29-MAR-2001; 2001US-0279882P.
29-MAR-2001; 2001US-0279884P.
30-MAR-2001; 2001US-0280147P.
11-APR-2001; 2001US-0282992P.
11-APR-2001; 2001US-0283083P.
20-APR-2001; 2001US-0285133P.
23-APR-2001; 2001US-0285749P.
03-MAY-2001; 2001US-0288327P.
03-MAY-2001; 2001US-0288504P.
29-MAY-2001; 2001US-0294047P.
30-MAY-2001; 2001US-0294473P.
08-JUN-2001; 2001US-0296964P.
18-JUN-2001; 2001US-0298959P.
19-JUN-2001; 2001US-0299324P.
13-AUG-2001; 2001US-0312020P.
16-AUG-2001; 2001US-0312889P.
21-AUG-2001; 2001US-0312908P.
28-AUG-2001; 2001US-0313390P.
31-AUG-2001; 2001US-0315470P.
07-SEP-2001; 2001US-0316447P.
07-SEP-2001; 2001US-0318115P.
12-SEP-2001; 2001US-0318118P.
19-SEP-2001; 2001US-0318740P.
18-OCT-2001; 2001US-0323379P.
18-OCT-2001; 2001US-0330245P.
14-NOV-2001; 2001US-0330308P.
14-NOV-2001; 2001US-0332701P.
(CURA-) CURAGEN CORP.
Tchernev VT, Spytek KA, Zerhusen BD, Patturajan M, Shimkets RA;
Li L, Gangolli EA, Padigaru M, Anderson DW, Rastelli L, Miller CE;
Gerlach VL, Taupier RJ, Gusev VY, Colman SD, Wolenc AR, Pena CE;
Furtak K, Grosse WM, Alsobrook JP, Lepley DM, Rieger DK, Burgess CE;
WPI; 2002-706998/76.
New NOVX polypeptides and nucleic acids, useful for preventing or
treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
pharmacogenomics.
Disclosure; SEQ ID NO 586; 1498pp; English.
This invention relates to a novel nucleic acids, and encoded polypeptides
thereof, which have properties related to the stimulation of biochemical
or physiological responses in a cell, tissue, organ or organism.
Specifically, it refers to the use of biologically active fragments for
diagnostic and prognostic assays and furthermore in the treatment of
diverse pathological conditions. The present invention describes novel
human and murine NOVX proteins, as well as methods to modulate their
expression using antisense oligos, ribozymes and peptide nucleic acids.
The NOVX polypeptides, polynucleotides and antibodies are useful in
treating or preventing NOVX-associated disorders, e.g. cardiomyopathy,
atherosclerosis, cancer and diabetes. Furthermore, they may be used in

CC treating or preventing diseases such as inflammation, autoimmune
CC disorders, allergies, blood disorders, acquired immunodeficiency syndrome
CC (AIDS), obesity, asthma, immunoglobulin (Ig)A nephropathy, cirrhosis,
CC arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy
CC and epilepsy. Accordingly, these molecules have many activities including
CC cytostatic, cardiant, antiinflammatory, immunosuppressive, antiallergic,
CC haemostatic, anti-HIV, antidiabetic, antiarteriosclerotic, anorectic,
CC antiasthmatic, nephrotropic, antiarthritic, hepatotropic,
CC neuroprotective, nootropic, antibacterial, viricide, antiparasitic,
CC relaxant and anticonvulsant. In addition, they are useful in screening
CC assays to identify small molecules that modulate or inhibit, for example,
CC neurogenesis, wound healing and angiogenesis. The nucleic acids are also
CC used as in chromosome mapping, tissue typing, preventive medicine and
CC pharmacogenomics. This polypeptide is a homologue of a human NOVX protein
CC of the invention.
XX
SQ Sequence 302 AA;
Query Match 29.3%; Score 449; DB 5; Length 302;
Best Local Similarity 36.2%; Pred. No. 5.4e-38;
Matches 96; Conservative 52; Mismatches 99; Indels 18; Gaps 3;
QY 36 EEIANFPVRPSDVITVTPKSGTSLQEVVLVSQADPDEIGLNNIDEQLPVLVYPOP- 94
Db 36 DKIMNFQAKPDDLLISTYPKAGTTWTQEIVELIQNEGDEKSKRAPHQRPFFLEWKIPS 95
QY 95 ---GLDIIKELTSPRIKSHLPYRFLPSDLHNGDSKVITYMARNPKDLVVSYQFHSRLT 151
Db 96 LSGLEQAHAHMPSPRIKTLHPFLLPSPSLLKKNCKIIVARNPKDMVSYVHFQRMNKA 155
QY 152 MSYRGTFQFCRRFMNDKLGYSWFHVEHQRMSNVLFKYEDMHRDLVTVVEQLA 211
Db 156 LPAPGTWEEYFETFLAGKVCWGSWHEHVKGWWEAKDHRILYLFYEDMKCKPKHBIQKLA 215
QY 212 RFLGVSCDKAOLEALTEHCQLV---DOCCNAEALP-----VGRGVLGWLKDF 257
Db 216 EFIGKLDKVDKLVHYTSFDMVKQNPANYSSIPAEIMDHSISFFMRKGAVGDKKHIF 275
QY 258 TVSMNEKFDLVYKQMGKCDLTFDF 282
Db 276 TVAQNERFEDYKKQMTDRLTFHF 300
RESULT 42
ADF76915
ID ADF76915 standard; protein; 302 AA.
XX
AC ADF76915;
XX
DT 26-FEB-2004 (first entry)
XX
DE Novel human secreted and transmembrane protein SeqID 590.
XX
KW human; PRO; membrane bound protein; membrane bound receptor;
KW cell proliferation; cell migration; cell differentiation;
KW mitogenic factor; survival factor; cytotoxic factor;
KW differentiation factor; neuropeptide; hormone; cell receptor;
KW receptor-ligand interaction; cytostatic; chondrocyte; tumour.
XX
OS Homo sapiens.
XX
PN WO2003072035-A2.
XX
PD 04-SEP-2003.
XX
PF 21-FEB-2003; 2003WO-US005241.
XX
PR 22-FEB-2002; 2002US-0359461P.
XX
PA (GETH) GENENTECH INC.
XX
PI Bodary SC, Clark H, Hunte B, Jackman JK, Schoenfeld JR;
PI Williams PM, Wood WI, Wu TD;

XX
DT 18-NOV-2004 (first entry)
XX Human diagnostic and therapeutic pprotein SEQ ID NO:4252.
XX
XX gene therapy; human diagnostic and therapeutic polynucleotide; dithp.
XX
XX Homo sapiens.
OS

XX PN WO2004023973-A2.
XX PD 25-MAR-2004.
XX PF 12-SEP-2003; 2003WO-US028227.
XX PR 12-SEP-2002; 2002US-0410259P.
XX PR 12-SEP-2002; 2002US-0410260P.
XX PA (INCY-) INCYTE CORP.
XX PI Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
PI Harthehorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
PI Petalca CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirtson ES;
PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
PI Patury S, Shi X, Suarez CJ;
XX WPI; 2004-329368/30.
DR N-PSDB; ACN42655.
XX PT New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT in diagnosing a condition, disease or disorder associated with human
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT in gene mapping.
XX PS Claim 27; Page; 190pp; English.
XX CC The invention relates to novel diagnostic and therapeutic polynucleotides
CC selected from one of the 2722 sequences defined in the specification. A
CC polynucleotide of the invention may have a use in gene therapy. The human
CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
CC used to diagnose a particular condition, disease or disorder associated
CC with human molecules, e.g. cell proliferative disorders,
CC autoimmune/inflammatory disorder, developmental disorder, endocrine
CC disorder, neurological disorders, gastrointestinal disorders, or
CC infections caused by virus, bacteria, fungi or parasite. The dithp
CC molecules may also be used in genetic mapping, in identifying individuals
CC from minute biological samples, in detecting single nucleotide
CC polymorphisms, as molecular weight markers, and for somatic or germline
CC gene therapy. The present sequence represents a dithp protein of the
CC invention. Note: The sequence data for this patent is not represented in
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX SQ Sequence 327 AA;
Query Match 29.2%; Score 447; DB 8; Length 327;
Best Local Similarity 32.5%; Pred. No. 9.8e-38;
Matches 101; Conservative 51; Mismatches 101; Indels 58; Gaps 5;
QY 24 GVLPPFCRGMKEETANPRPSDVWIVTPKSGTSLQEVVLYVSQCADPDEIGLMNID 83
DB 17 GVPLIKYFAEALGPQSQFARDDLLINTPKSGTTWSQLDMYQGGDLKCNRAPIY 76
QY 84 EQLPVLEYQPQ---GLDIIKELTSPRLIKSHLPYRFLPSPDLHNGDSKVIYMARPKDLV 139
DB 77 VRVPFLEVNDQPEPSGLETLDKTPPRLIKSHLPALPQLTLDQKVVVYARNPKDVA 136
QY 140 VSYQFHSRLTMSVGTTFQFCRRFMDKLG----- 171
DB 137 VSYVHFHREMAKHPPEFGTWSFLEKFMAGEGLDWRKEGVKPRGGYVNVQPCVGAQPL 196
QY 172 ---YGSFHFVQEFWEHRMDSNVLFKLYEDMHRDLVTMVEQLARFLGVSCDKQALELT 227
DB 197 LVXVSGSWYQVQEWELSRTHPVLYLFYEDMKENPKRIQKILFVGRSLPEETMDPMV 256
QY 228 EHC-----HQLVDOCCNAEALP-VGRGRVGLMKDIFTVSNKEFDLVY 269
DB 257 QHTSPKMKKNPMTNTYTTVPQLMDHSIS----PFMRKMGMDGDKTTFVAGNERFDADY 312

QY 270 KQKMGKCDLTF 280
DB 313 AERWAGCSLSF 323
RESULT 49
ABM84004
ID ABM84004 standard; protein; 327 AA.
XX AC ABM84004;
XX DT 18-NOV-2004 (first entry)
XX DE Human diagnostic and therapeutic pprotein SEQ ID NO:4253.
XX KW gene therapy; human diagnostic and therapeutic polynucleotide; dithp.
XX OS Homo sapiens.
XX PN WO2004023973-A2.
XX PD 25-MAR-2004.
XX PF 12-SEP-2003; 2003WO-US028227.
XX PR 12-SEP-2002; 2002US-0410259P.
XX PR 12-SEP-2002; 2002US-0410260P.
XX PA (INCY-) INCYTE CORP.
XX PI Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
PI Harthehorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
PI Petalca CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirtson ES;
PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
PI Patury S, Shi X, Suarez CJ;
XX WPI; 2004-329368/30.
DR N-PSDB; ACN42656.
XX PT New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT in diagnosing a condition, disease or disorder associated with human
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT in gene mapping.
XX PS Claim 27; Page; 190pp; English.
XX CC The invention relates to novel diagnostic and therapeutic polynucleotides
CC selected from one of the 2722 sequences defined in the specification. A
CC polynucleotide of the invention may have a use in gene therapy. The human
CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
CC used to diagnose a particular condition, disease or disorder associated
CC with human molecules, e.g. cell proliferative disorders,
CC autoimmune/inflammatory disorder, developmental disorder, endocrine
CC disorder, neurological disorders, gastrointestinal disorders, or
CC infections caused by virus, bacteria, fungi or parasite. The dithp
CC molecules may also be used in genetic mapping, in identifying individuals
CC from minute biological samples, in detecting single nucleotide
CC polymorphisms, as molecular weight markers, and for somatic or germline
CC gene therapy. The present sequence represents a dithp protein of the
CC invention. Note: The sequence data for this patent is not represented in
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX SQ Sequence 327 AA;
Query Match 29.2%; Score 447; DB 8; Length 327;
Best Local Similarity 32.5%; Pred. No. 9.8e-38;
Matches 101; Conservative 51; Mismatches 101; Indels 58; Gaps 5;


```
XX SQ Sequence 269 AA;
Query Match 29.0%; Score 444; DB 5; Length 269;
Best Local Similarity 40.6%; Pred. NO. 1.5e-37;
Matches 104; Conservative 39; Mismatches 83; Indels 30; Gaps 7;

Qy 41 FVPSDVIWYTPKSGTSLQEVVLYVSQAD-----PDEIGLMNIDEQLPVLEYPPQG 95
Db 20 FOARDDVLIAGPKSGTWTQEIULSHPNVGDFFPSPSDPLLFNR-----PWLEYPK-G 73

Qy 96 LDIKEL----TSPLIKSHLPYRPLPSDLHNGDSKVYIMARNPKDLVVSYQFHRSLRT 151
Db 74 EDWYETLAPMSSPLIKTHLPLELLPKSFLSSKAKIIVLNPKNDAVSYHFSRSHKD 133

Qy 152 M-SYRGTFQECRRPMNDKLVGSGFHFVQEFWEHRMDSNVFLKYEDMHRDLVTMVQOL 210
Db 134 LPADPGTFEEFLAFLNGKLVGSGFYFDHVLGWELRPEPQVFLFDYEDLKEDPAGEIKKI 193

Qy 211 ARFLGVSCDKAQLAETEHCH---QLVQCCNAEALPVG-----BGRVGLWKDI 256
Db 194 AEFGLPLSEBELDKLHDHSPFLMKJNPLGNYETLCIGKSKGRKSPFMRKGLVGDWKNY 253

Qy 257 FTVSMNEKFDLVYKOK 272
Db 254 FTPEQNEKFDVKEK 269

RESULT 51
AD117331
ID AD117331 standard; protein; 269 AA.
XX AC AD117331;
XX DT 15-APR-2004 (first entry)
XX DE Polypeptide homologous to a human NOVX domain SeqID 867.
XX KW NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes; inflammation;
XX autoimmunity disorder; allergy; blood disorder;
XX acquired immunodeficiency syndrome; AIDS; obesity; asthma;
XX immunoglobulin (Ig)A nephropathy; cirrhosis; arthritis;
XX Alzheimer's disease; infection; str.
XX OS Unidentified.
XX PN WO200268649-A2.
XX PD 06-SEP-2002.
XX PF 31-JAN-2002; 2002WO-US002785.
XX PR 31-JAN-2001; 2001US-0265395P.
XX PR 31-JAN-2001; 2001US-0265412P.
XX PR 31-JAN-2001; 2001US-0265514P.
XX PR 31-JAN-2001; 2001US-0265517P.
XX PR 02-FEB-2001; 2001US-0266408P.
XX PR 05-FEB-2001; 2001US-0266767P.
XX PR 07-FEB-2001; 2001US-0266975P.
XX PR 08-FEB-2001; 2001US-0267057P.
XX PR 08-FEB-2001; 2001US-0267459P.
XX PR 09-FEB-2001; 2001US-0267823P.
XX PR 15-FEB-2001; 2001US-0268974P.
XX PR 26-FEB-2001; 2001US-0271664P.
XX PR 27-FEB-2001; 2001US-0271839P.
XX PR 27-FEB-2001; 2001US-0271855P.
XX PR 02-MAR-2001; 2001US-0272789P.
XX PR 02-MAR-2001; 2001US-0273046P.
XX PR 14-MAR-2001; 2001US-0275925P.
XX PR 14-MAR-2001; 2001US-0275947P.
XX PR 14-MAR-2001; 2001US-0275950P.
XX PR 14-MAR-2001; 2001US-0275989P.
XX PR 15-MAR-2001; 2001US-0276448P.
PR 15-MAR-2001; 2001US-0276450P.
PR 16-MAR-2001; 2001US-0276397P.
PR 16-MAR-2001; 2001US-0276768P.
PR 20-MAR-2001; 2001US-0276768P.
PR 26-MAR-2001; 2001US-0278652P.
PR 26-MAR-2001; 2001US-0278775P.
PR 29-MAR-2001; 2001US-0278778P.
PR 29-MAR-2001; 2001US-0279882P.
PR 29-MAR-2001; 2001US-0279884P.
PR 30-MAR-2001; 2001US-0280147P.
PR 11-APR-2001; 2001US-0282992P.
PR 11-APR-2001; 2001US-0283083P.
PR 20-APR-2001; 2001US-0285133P.
PR 23-APR-2001; 2001US-0285749P.
PR 03-MAY-2001; 2001US-0288327P.
PR 03-MAY-2001; 2001US-0288504P.
PR 29-MAY-2001; 2001US-0294047P.
PR 30-MAY-2001; 2001US-0294473P.
PR 08-JUN-2001; 2001US-0296964P.
PR 18-JUN-2001; 2001US-0298959P.
PR 19-JUN-2001; 2001US-0299324P.
PR 13-AUG-2001; 2001US-0312020P.
PR 16-AUG-2001; 2001US-0312889P.
PR 16-AUG-2001; 2001US-0312908P.
PR 21-AUG-2001; 2001US-0313390P.
PR 28-AUG-2001; 2001US-0315470P.
PR 31-AUG-2001; 2001US-0316447P.
PR 07-SEP-2001; 2001US-0318115P.
PR 07-SEP-2001; 2001US-0318118P.
PR 12-SEP-2001; 2001US-0318740P.
PR 19-SEP-2001; 2001US-0323379P.
PR 18-OCT-2001; 2001US-0330245P.
PR 18-OCT-2001; 2001US-0330308P.
PR 14-NOV-2001; 2001US-0332701P.
XX PA (CURA-) CURAGEN CORP.
XX Tchernev VT, Spytek KA, Zerhusen BD, Patturajan M, Shimkets RA;
XX Li L, Gangolli EA, Padigaru M, Anderson DW, Rastelli L, Miller CE;
XX Gerlach VR, Taupier RJ, Gusev VY, Colman SD, Wolenc AR, Pena CEA;
XX Furtak K, Grosse MW, Alsobrook JP, Lepley DM, Rieger DK, Burgess CE;
XX WPI; 2002-706998/76.
XX New NOVX polypeptides and nucleic acids, useful for preventing or
XX treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
XX atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
XX pharmacogenomics.
XX PS Disclosure; SEQ ID NO 867; 1498pp; English.
XX CC This invention relates to a novel nucleic acids, and encoded polypeptides
XX thereof, which have properties related to the stimulation of biochemical
XX or physiological responses in a cell, tissue, organ or organism.
XX Specifically, it refers to the use of biologically active fragments for
XX diagnostic and prognostic assays and furthermore in the treatment of
XX diverse pathological conditions. The present invention describes novel
XX human and murine NOVX proteins, as well as methods to modulate their
XX expression using antisense oligos, ribozymes and peptide nucleic acids.
XX The NOVX polypeptides, polynucleotides and antibodies are useful in
XX treating or preventing NOVX-associated disorders, e.g. cardiomyopathy,
XX atherosclerosis, cancer and diabetes. Furthermore, they may be used in
XX treating or preventing diseases such as inflammation, autoimmune
XX disorders, allergies, blood disorders, acquired immunodeficiency syndrome
XX (AIDS), obesity, asthma, immunoglobulin (Ig)A nephropathy, cirrhosis,
XX arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy
XX and epilepsy. Accordingly, these molecules have many activities including
XX cytostatic, cardiant, antiinflammatory, immunosuppressive, antiallergic,
XX haemostatic, anti-HIV, antidiabetic, antiarteriosclerotic, anorectic,
XX antiasthmatic, nephrotropic, antiarthritic, hepatotropic,
XX neuroprotective, neurotropic, antibacterial, virucide, antiparasitic,
XX relaxant and anticonvulsant. In addition, they are useful in screening
XX assays to identify small molecules that modulate or inhibit, for example,
XX neurogenesis, wound healing and angiogenesis. The nucleic acids are also
```

CC used as in chromosome mapping, tissue typing, preventive medicine and
CC pharmacogenomics. This polypeptide is a protein fragment that is
CC homologous to a domain of a human NOVX protein of the invention.
XX
SQ Sequence 269 AA;
Query Match 29.0%; Score 444; DB 5; Length 269;
Best Local Similarity 40.6%; Pred. No. 1.5e-37;
Matches 104; Conservative 39; Mismatches 83; Indels 30; Gaps 7;
QY 41 FVPRSDVIVTPYKSGTSLLOEVVYVLSQGD-----PDEIGLMNIDQLPVLVYPQPG 95
DB 20 FQARPDVLIAGVPGSGTTLQELSLHPNVGDFEPPSPDPLFRN-----PWLEVPK-G 73
QY 96 LDIIEKL-----TSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPDLVVSYYQFHSRLRT 151
DB 74 EDWYETLKPMPSPRLIKTHLPKSLFLSKAKIIVLRNPKDVAVSYHFSRSHKD 133
QY 152 M-SYRGTFQFCRRFMNDKLGVSFHFVQEFWEHRMDSNVLPFLKVEDMHRDLVTMVEQL 210
DB 134 LPADPGTFEELEAFNGHVKVLYGSPFDHVGWELRPEPQVLFLEDYEDJEDPAGEIKKI 193
QY 211 ARFLGVSCDKAQLALETCH-----QLVDOCCNAEALPVG-----RGRVGLWKDI 256
DB 194 ABFLGLPSEELDKLHSSFFLMKLNPLSNVETLCGSKGRKSPFMRKGLVGDWKNY 253
QY 257 FTVSMNEKPDLYVYKOK 272
DB 254 FTPEQNEKPKVKEK 269
RESULT 52
ADD12573
ID ADD12573 standard; protein; 310 AA.
XX
AC ADD12573;
XX
DT 01-JAN-2004 (first entry)
DE Human ENZM-33 protein SEQ ID NO:33.
XX
KW human; enzyme; ENZM; cytostatic; antiarteriosclerotic; anti-HIV;
KW antiallergic; cerebroprotective; antiparkinsonian; anticonvulsant;
KW neurologic; neuroprotective; antiinflammatory; ophthalmological;
KW antithyroid; antirheumatic; antibacterial; virucide; protozoacide;
KW antiparasitic; fungicide; anorectic; cardiant; hypotensive;
KW antinfertility; hepatotropic; gene therapy; autoimmune disorder;
KW inflammatory disorder; AIDS; allergy; atopic dermatitis; arthritis;
KW thyroiditis; infection; metabolic disorder; obesity;
KW reproductive disorder; infertility; neurological disorder;
KW Parkinson's disease; Alzheimer's disease; stroke; epilepsy;
KW cardiovascular disorder; myocardial infarction; hypertension;
KW eye disorder; cell proliferative disease; cancer; atherosclerosis;
KW hepatitis.
XX
OS Homo sapiens.
XX
PN WO2003072729-A2.
XX
PD 04-SEP-2003.
XX
PF 21-FEB-2003; 2003WO-US005478.
XX
PR 22-FEB-2002; 2002US-0359513P.
PR 19-MAR-2002; 2002US-0365795P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
PI Sprague WW, Kable AE, Lee S, Griffin JA, Becha SD, Lee EA;
PI Hafalia AJA, Khare R, Emerling BM, Marquis JP, Ramkumar J;
PI Elliott VS, Richardson TW, Baughn MR, Jin P, Chien D, Hawkins PR;
PI Wilton AD, Chawla NK, Tran UK, Lee SY, Zebajradian Y, Jiang X;
PI Jackson AA, Bhatia UG, Burrill JD, Blake JJ, Ho A, Zheng W;

XX WPI: 2003-731608/69.
DR N-PSDB; ADD12630.
XX
PT New human enzymes and polynucleotides, useful for diagnosing, preventing
PT or treating diseases or conditions associated with aberrant enzyme
PT expression, e.g. cancer, hepatitis, AIDS, atherosclerosis, obesity,
PT infection or stroke.
XX
PS Claim 1; SEQ ID NO 33; 426pp; English.
XX
CC The present sequence represents a human enzyme designated ENZM-33. Human
CC enzymes of the present invention have cytostatic, antiarteriosclerotic,
CC anti-HIV, antiallergic, cerebroprotective, antiparkinsonian,
CC anticonvulsant, neurologic, neuroprotective, antiinflammatory,
CC ophthalmological, antithyroid, antirheumatic, antibacterial, virucide,
CC protozoacide, antiparasitic, fungicide, anorectic, cardiant, hypotensive,
CC antinfertility and hepatotropic activities, and can be used in gene
CC therapy. The human enzymes and polynucleotides encoding them can be used
CC in diagnosing, preventing or treating diseases or conditions associated
CC with the decreased expression or overexpression of the enzymes, such as
CC autoimmune/inflammatory disorders (AIDS, allergies, atopic dermatitis,
CC arthritis or thyroiditis), infections (e.g. bacterial, viral, parasitic,
CC protozoal or fungal), metabolic disorders (e.g. obesity), reproductive
CC disorders (e.g. infertility), neurological disorders (Parkinson's
CC disease, Alzheimer's disease, stroke, epilepsy), cardiovascular disorders
CC (e.g. myocardial infarction or hypertension), eye disorders, or cell
CC proliferative diseases (e.g. cancer, atherosclerosis, hepatitis). They
CC are also useful in assessing the effects of exogenous compounds on the
CC expression of nucleic acid and amino acid sequences of the enzymes.
XX
SQ Sequence 310 AA;
Query Match 28.9%; Score 442; DB 7; Length 310;
Best Local Similarity 31.5%; Pred. No. 3e-37;
Matches 94; Conservative 61; Mismatches 105; Indels 38; Gaps 4;
QY 18 KYFPHGVRLPPFCRGKMEETANFVRPSDVWIVTPYKSGTSLLOEVVYVLSQGDPPDEI 77
DB 12 KLRKEVEGTLQPATVDNWSQISQFPAKDDLLICTYKAGTWTWIEQDMTEQNDVEKC 71
QY 78 GLMNIDEOLPVLVY---POP-----GLDIIKELTSPRIKSHLPVRLFPS 119
DB 72 QRALIQHRHPFIEWARPPQPSVULVRCFLSYFGVGEKAKAMPSPRIKTHLSTQLLPP 131
QY 120 DLHNGDSKVIYMARNPDLVVSYYQFHSRLRTMSYRGTFQFCRRFMNDKLGYSWFHIV 179
DB 132 SFWENNCKFLVARNNAKDCMVSYHHFQRMNHLPPDGTWEEYFETTINGKVMGWSFHDV 191
QY 180 QEFWEHRMDSNVLPFLKVEDMHRDLVTMVEQLARFLGVSCDKAQLALETCH----- 229
DB 192 KGWEMKDRHQILFLFYEDIKRDPKHEIRKVMQFMGKKVDETVLDKIVQETSFERMKENP 251
QY 230 -----CHQLVDOCCNAEALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
DB 252 MTRNSTVSKSILDQSISSF---MRKGTVDGNHFTVAQNERFDEIYRRKMEGTSINF 306
RESULT 53
ADA10983
ID ADA10983 standard; protein; 296 AA.
XX
AC ADA10983;
XX
DT 06-NOV-2003 (first entry)
XX
DE Human cDNA differentially expressed in colon cancer #68 product.
KW differential expression; colon cancer; cancer; human.
XX
OS Homo sapiens.
PN US2002160382-A1.

XX Human; oestrogen sulphotransferase; estrogen sulphotransferase; liver EST;
KW biotransformation.
XX
XX Homo sapiens.
XX
XX US5714594-A.
XX
XX 03-FEB-1998.
XX
XX 18-OCT-1994; 94US-00325562.
XX
XX 18-OCT-1994; 94US-00325562.
XX
XX (MAYO-) MAYO FOUNDATION.
XX
XX Wood TC, Weinshilboum RM, Aksoy IA;
XX
XX WPI; 1998-144284/13.
XX
XX N-PSDB; AAV12466.
XX
XX DNA encoding human oestrogen sulphotransferase - useful for research into
XX oestrogen biotransformation.
XX
XX Claim 1; Col 21-24; 25pp; English.
XX
XX The present sequence represents a human oestrogen sulphotransferase (EST)
XX protein. The isolation and expression of cDNA which encodes human liver
XX EST enables the biotransformation of oestrogens to be studied
XX
XX Sequence 294 AA;
XX
XX Query Match 28.4%; Score 434; DB 2; Length 294;
XX Best Local Similarity 33.3%; Pred. No. 1.9e-36;
XX Matches 95; Conservative 56; Mismatches 116; Indels 18; Gaps 3;
XX
XX 14 EFESKYFEFHGVRLPPFCRCRMEETANFPVRPSDVWIVTPKSGTSLLOEVVYLVSQAD 73
XX 6 DYKFEFVHGILMYKDFVKYWDNVEAFQARDDLVVATYPKSGTWTWSEIVYMYIKRGD 65
XX
XX 74 PDEIGLNMIDQLPVLEYPQP----GLDIIKELTSPLRIKSHLPYRFLPSDLHNGDSKVI 129
XX 66 VEKCKEDVFNRIPELECRKENLNGVKQLDEMNSPRIVKTHLPPELLPASFWKDCCKII 125
XX
XX 130 YMARNPXDLVVSYYQFHRSLRMTSYRGTFQFCRRFMDKLGYSWFEHVFWEHRMDS 189
XX 126 YLCRNADKDVAVSFYFFFLMVAGHPNPGSFPEFVEKFMQGVPGYSGWYKHKVSWKEGKSP 185
XX
XX 190 NVFLKYMEDHRDLVTWVEQLARFLGVSCDKAQLAALTEHCH-----QLVDQ 236
XX 186 RVLFYFEDLKBEDIRKEVIKLHFLEKPSSELVDRIIHHTSFQEMKNPNSTNYTTLPDE 245
XX
XX 237 CCNAEALP-VGRGRVGLWKIDIFTVSMNEKFDLVYKQKMGKCDLTF 280
XX 246 IMNQKLSFPMKGIIGDKNHFTVALNEKFDKHYEQQKESTLKF 290
XX
XX RESULT 56
XX ADD48896
XX ID ADD48896 standard; protein; 294 AA.
XX
XX AC ADD48896;
XX
XX 02-DEC-2004 (revised)
XX 29-JAN-2004 (first entry)
XX
XX Human Protein P49888, SEQ ID NO 14607.
XX
XX Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
XX spared nerve injury; SNI; Chung.
XX
XX Homo sapiens.
OS

OS Unidentified.
XX WO2003016475-A2.
XX
XX 27-FEB-2003.
XX
XX 14-AUG-2002; 2002WO-US025765.
XX
XX 14-AUG-2001; 2001US-0312147P.
XX
XX 01-NOV-2001; 2001US-0346382P.
XX
XX 26-NOV-2001; 2001US-0333347P.
XX
XX (GEO) GEN HOSPITAL CORP.
XX
XX (FARB) BAYER AG.
XX
XX Woolf C, D'urso D, Befort K, Costigan M;
XX
XX WPI; 2003-268312/26.
XX
XX GENBANK; P49888.
XX
XX New composition comprising two or more isolated polypeptides, useful for
XX preparing a medicament for treating pain in an animal.
XX
XX Example 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
XX or human polynucleotides or a polynucleotide which represents a fragment,
XX derivative or allelic variation of the nucleic acid sequence. Also
XX claimed are a vector comprising the novel polynucleotide, a host cell
XX comprising the vector, a method for identifying a nucleotide sequence
XX which is differentially regulated in an animal subjected to pain and a
XX kit to perform the method, an array, a method for identifying an agent
XX that increases or decreases the expression of the polynucleotide sequence
XX that is differentially expressed in neuronal tissue of a first animal
XX subjected to pain, a method for identifying a compound which regulates
XX the expression of a polynucleotide sequence which is differentially
XX expressed in an animal subjected to pain, a method for identifying a
XX compound that regulates the activity of one or more of the
XX polynucleotides, a method for producing a pharmaceutical composition, a
XX method for identifying a compound or small molecule that regulates the
XX activity in an animal of one or more of the polypeptides given in the
XX specification, a method for identifying a compound useful in treating
XX pain and a pharmaceutical composition comprising the one or more
XX polypeptides or their antibodies. The polynucleotide or the compound that
XX modulates its activity is useful for preparing a medicament for treating
XX pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
XX injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
XX therapy). The sequence presented is a human protein (described in Table 3
XX of the specification) which is differentially expressed during pain.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic form directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 294 AA;
XX
XX Query Match 28.4%; Score 434; DB 7; Length 294;
XX Best Local Similarity 33.3%; Pred. No. 1.9e-36;
XX Matches 95; Conservative 56; Mismatches 116; Indels 18; Gaps 3;
XX
XX 14 EFESKYFEFHGVRLPPFCRCRMEETANFPVRPSDVWIVTPKSGTSLLOEVVYLVSQAD 73
XX 6 DYKFEFVHGILMYKDFVKYWDNVEAFQARDDLVVATYPKSGTWTWSEIVYMYIKRGD 65
XX
XX 74 PDEIGLNMIDQLPVLEYPQP----GLDIIKELTSPLRIKSHLPYRFLPSDLHNGDSKVI 129
XX 66 VEKCKEDVFNRIPELECRKENLNGVKQLDEMNSPRIVKTHLPPELLPASFWKDCCKII 125
XX
XX 130 YMARNPXDLVVSYYQFHRSLRMTSYRGTFQFCRRFMDKLGYSWFEHVFWEHRMDS 189
XX 126 YLCRNADKDVAVSFYFFFLMVAGHPNPGSFPEFVEKFMQGVPGYSGWYKHKVSWKEGKSP 185
XX
XX 190 NVFLKYMEDHRDLVTWVEQLARFLGVSCDKAQLAALTEHCH-----QLVDQ 236
XX
XX


```
XX Claim 25; Fig 2; 103pp; English.
PS The invention relates to a novel human sulfotransferase polypeptide. The
XX sulfotransferase has cytosolic activity. The polynucleotide of the
CC invention may have a use in gene therapy. The polypeptide and
CC polynucleotide of the invention are useful in preventing, ameliorating,
CC or correcting diseases including cancer and chronic obstructive pulmonary
CC disease (COPD). The sequence represents a human sulfotransferase of the
XX invention
XX Sequence 304 AA;
SQ
Query Match 27.6%; Score 422; DB 5; Length 304;
Best Local Similarity 32.3%; Pred. No. 3.7e-35;
Matches 97; Conservative 59; Mismatches 124; Indels 20; Gaps 5;
QY 1 MAESAETPSTGPEE-SKYFEFHGVRLLPPFCRGKMEETANFPVRPSDVWIVTPKSGTS 59
DB 1 MAKIEKNAPTWEKKPELFINMEVDGVTLLLSKEWKEKVCNFOAKPDDLILATYPKSGTT 60
QY 60 LQEVVYLVSGADPDEIGLMNIDQLPVL-YP---QGLDIIKELTSPRLIKSHLPY 114
DB 61 WHHEILMDLNDGDVEKCKRAQTLDRHAFLELKFPHKEKPDLEFVLEMSSPQIKTHLPS 120
QY 115 RFLPSDLHNGDSKVIYMARNPDKLVVYVQFHRSLRTMSYRGTFQFCRRFMDKLGYGS 174
DB 121 HLIPPSIWKENCKIYVARNPKDCLVSYHFRHMASFMPDPQNLEEFYEKFMGKVGGS 180
QY 175 WFEHQVEFWEHRMDSNVLPLKYEDMHRDLVTMVEQLARFLGVSCDAQLEALTEHQVLV 234
DB 181 WFDVKGWAAKMDHRILYLFYEDIKDKPKREIKLFEKDIIEEILNKIIYHTSDV 240
QY 235 ---DQCNAEALPVG-----RGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTF 280
DB 241 MKQNPMTNVTTLPTSIMDHSIPFWRKGMPGDKWKNYFTVAQNEEFDKDYQKMGAGTLTF 300
RESULT 60
ID ADD48894 standard; protein; 295 AA.
XX ADD48894;
XX ADD48894;
XX 02-DEC-2004 (revised)
DT 29-JAN-2004 (first entry)
XX Rat Protein P52844, SEQ ID NO 14605.
XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
XX Chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX Rattus norvegicus.
OS Unidentified.
XX WO2003016475-A2.
XX 27-FEB-2003.
XX 14-AUG-2002; 2002WO-US025765.
XX 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX (GEHO ) GEN HOSPITAL CORP.
PA (FARB ) BAYER AG.
XX Woolf C, D'urso D, Befort K, Costigan M;
XX WPI; 2003-268312/26.
DR GENBANK; P52844.
XX
```

New composition comprising two or more isolated polypeptides, useful for preparing a medicament for treating pain in an animal.

Example 1; Page; 1017pp; English.

The invention discloses a composition comprising two or more isolated rat or human polynucleotides or a polynucleotide which represents a fragment, derived or allelic variation of the nucleic acid sequence. Also claimed are a vector comprising the novel polynucleotide, a host cell comprising the vector, a method for identifying a nucleotide sequence which is differentially regulated in an animal subjected to pain and a kit to perform the method, an array, a method for identifying an agent that increases or decreases the expression of the polynucleotide sequence that is differentially expressed in neuronal tissue of a first animal subjected to pain, a method for identifying a compound which regulates the expression of a polynucleotide sequence which is differentially expressed in an animal subjected to pain, a method for identifying a compound that regulates the activity of one or more of the polynucleotides, a method for producing a pharmaceutical composition, a method for identifying a compound or small molecule that regulates the activity in an animal of one or more of the polypeptides given in the specification, a method for identifying a compound useful in treating pain and a pharmaceutical composition comprising the one or more polypeptides or their antibodies. The polynucleotide or the compound that modulates its activity is useful for preparing a medicament for treating pain (e.g. spinal segmental nerve injury (Chung), chronic constriction injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene therapy). The sequence presented is a rat protein (described in Table 3 of the specification) which is differentially expressed during pain. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic form directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 295 AA;

Query Match 27.5%; Score 421; DB 7; Length 295;
Best Local Similarity 31.9%; Pred. No. 4.5e-35;
Matches 91; Conservative 60; Mismatches 116; Indels 18; Gaps 3;

QY 14 EFESKYFEFHGVRLLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSLQEVVYLVSGAD 73
DB 7 EYEVFGDFHGLMDKRFYKWEDETFLARPDDLIVTPKSGSTWISIVDMYKRGD 66
QY 74 PDEIGLMNIDQLPVLVYQPQ---GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVI 129
DB 67 VEKCKEDALFNRIIDLECRNEDLINGIKQLKEKSPRIVKTHLPAKLPAFWEKNCXII 126
QY 130 YMARNPDKLVVSYVQFHRSLRTMSYRGTFQFCRRFMDKLGYGSWFHVOEFWEHRMDS 189
DB 127 YLCRNADVVVSYVYFFFLIMKSYNPKSFSEFVEKFMGQVPGSWYDHVKSWEKSKNS 186
QY 190 NVLFLKYEDMHRDLVTMVEQLARFLGVSCDAQLEALTEHC---HQLVDQCNAEALP-- 244
DB 187 RVLFMFYEDMKEDIRREVVKLIEFLERDPSAELVDRIIQTSTFOEMKKNPCTNYSMLPET 246
QY 245 -----VGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTF 280
DB 247 MIDLVSPFMRKGI VGDWRNHFPEALRFRFEEYQRHMKDCPVKF 291

RESULT 61
ABB81793
ID ABB81793 standard; protein; 304 AA.
XX ABB81793;
XX ABB81793;
DT 24-SEP-2002 (first entry)
XX Human sulfotransferase #3.
DE Human; sulfotransferase; cytosolic; gene therapy; cancer; COPD;
KW chronic obstructive pulmonary disease.
XX

PR 20-JUN-2000; 2000US-0212725P.
PR 03-JUL-2000; 2000US-00609816.
XX (PEKE) PE CORP NY.
PA
XX Woodage T, Wei M, Kodira C, Di Francesco V, Beasley EM;
XX WPI; 2001-616484/71.
XX N-PSDB; RAD19776.
XX
XX New polypeptides useful in labeling reagents with tagged sulfur and
PT identifying modulators of the protein comprises proteins related to
PT sulfotransferase drug-metabolizing enzyme subfamily.
XX
XX Claim 1; Fig 2; 83pp; English.
XX
CC The invention relates to sulfotransferase (ST) drug-metabolising enzyme
CC (DME) peptides, proteins and nucleic acid molecules encoding them.
CC Sequences of the invention are useful in labelling reagents with tagged
CC sulphur, for the development of human therapeutics, as query sequences to
CC perform a search against sequence databases, for identifying other family
CC members or related sequences, to raise antibodies or to elicit another
CC immune response, as a reagent in assays designed to quantitatively
CC determine levels of the protein in biological fluids, to metabolise
CC compounds for toxicity studies, and as markers for tissues in which the
CC corresponding protein is preferentially expressed. Polypeptides of the
CC invention is also useful in diagnosing a disease, predisposition to a
CC disease mediated by the peptide, in pharmacogenomic analysis and treating
CC a disorder characterised by absence of inappropriate or unwanted
CC expression of the protein. Polynucleotides of the invention are useful
CC for constructing recombinant vectors and host cells, expressing antigenic
CC portions of the proteins, in genotyping and mapping experiments, for
CC designing ribozymes, for monitoring the effectiveness of modulating
CC compounds on the expression or activity of the protein, in diagnostic
CC assays, antisense constructs to control protein expression and in gene
CC therapy. The present sequence is ST drug-metabolising protein 1 encoded
CC by DNA transcript 1
XX
SQ Sequence 304 AA;
Query Match 26.2%; Score 401; DB 4; Length 304;
Best Local Similarity 31.0%; Pred. No. 5.9e-33;
Matches 93; Conservative 62; Mismatches 125; Indels 20; Gaps 5;
QY 1 MAESAEPTSTGEFE-SKYFEFHGVLPPFCRGKMEETANFPVPSDWIIVTPKSGTS 59
DB 1 MAKIEKNAPTMEKKPELFNIMEVDGVPTLLISKWEWKCVCNFOAKPDDLLIATYPRSGTT 60
QY 60 LIQEVVYLVSQGADPEIGLANNIDQLPVL--YP---OPGLDIIKELTSPRLIKSHLPY 114
DB 61 WHEILDMLNDGDVEKCKRAQTLDRAFLKLPFKHKEKPDLEFVLEMSSPQLIKTHLPS 120
QY 115 RPLPSDLHNGDSKVITYMARNPKDLVSVYQFHRSLRTWSYRGTFOEFCRRFMNDKLGYS 174
DB 121 HLIPPSIWKENCKIVYVARNPKDLVSVYHFRMASFMPDPQNLBEFYEKFGSKGVGGS 180
QY 175 WFEHQVEWHRMDSNVLFKVEDMHRDLVTWVEQLARFLGVSCDKAQLALETECHQLV 234
DB 181 WFDHVKWMAAKDMHRILYLFYFKKPKHIEHKVLEFLEKTWSGDVINKIVHRTSFDV 240
QY 235 ---DOCCNAEALP-----VGRGRVGLWKDIFTVMNEKFDLVYKQMGKCDLTF 280
DB 241 MKDNPMANHTAVPAHFNHISKFMKGMFGDKKHNFTVALNENFDKHYEKWAGSTLNF 300
RESULT 69
AD117046
ID AD117046 standard; protein; 304 AA.
XX
AC AD117046;
XX
DT 15-APR-2004 (first entry)
XX

DE Human NOVX protein homologue SeqID 582.
XX human; NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes;
KW inflammation; autoimmune disorder; allergy; blood disorder;
KW acquired immunodeficiency syndrome; AIDS; obesity; asthma;
KW immunoglobulin (Ig)A nephropathy; cirrhosis; arthritis;
KW Alzheimer's disease; infection; str.
XX
OS Homo sapiens.
XX
XX WO2002068649-A2.
XX
PD 06-SEP-2002.
XX
PF 31-JAN-2002; 2002WO-US002785.
XX
PR 31-JAN-2001; 2001US-0265395P.
PR 31-JAN-2001; 2001US-0265412P.
PR 31-JAN-2001; 2001US-0265514P.
PR 31-JAN-2001; 2001US-0265517P.
PR 02-FEB-2001; 2001US-0266406P.
PR 05-FEB-2001; 2001US-0266767P.
PR 07-FEB-2001; 2001US-0266975P.
PR 07-FEB-2001; 2001US-0267057P.
PR 08-FEB-2001; 2001US-0267459P.
PR 09-FEB-2001; 2001US-0267823P.
PR 15-FEB-2001; 2001US-0268974P.
PR 26-FEB-2001; 2001US-0271664P.
PR 27-FEB-2001; 2001US-0271839P.
PR 27-FEB-2001; 2001US-0271855P.
PR 02-MAR-2001; 2001US-0272788P.
PR 02-MAR-2001; 2001US-0273046P.
PR 14-MAR-2001; 2001US-0275925P.
PR 14-MAR-2001; 2001US-0275947P.
PR 14-MAR-2001; 2001US-0275950P.
PR 14-MAR-2001; 2001US-0275989P.
PR 15-MAR-2001; 2001US-0276448P.
PR 15-MAR-2001; 2001US-0276450P.
PR 16-MAR-2001; 2001US-0276397P.
PR 16-MAR-2001; 2001US-0276768P.
PR 20-MAR-2001; 2001US-0278652P.
PR 26-MAR-2001; 2001US-0278775P.
PR 26-MAR-2001; 2001US-0278778P.
PR 29-MAR-2001; 2001US-0279882P.
PR 29-MAR-2001; 2001US-0279884P.
PR 30-MAR-2001; 2001US-0280147P.
PR 11-APR-2001; 2001US-0282992P.
PR 11-APR-2001; 2001US-0283083P.
PR 20-APR-2001; 2001US-0285133P.
PR 03-MAY-2001; 2001US-0285749P.
PR 23-APR-2001; 2001US-0288327P.
PR 03-MAY-2001; 2001US-0288504P.
PR 29-MAY-2001; 2001US-0294047P.
PR 30-MAY-2001; 2001US-0294473P.
PR 08-JUN-2001; 2001US-0296964P.
PR 18-JUN-2001; 2001US-0298959P.
PR 19-JUN-2001; 2001US-0299324P.
PR 13-AUG-2001; 2001US-0312020P.
PR 16-AUG-2001; 2001US-0312889P.
PR 21-AUG-2001; 2001US-0312908P.
PR 16-AUG-2001; 2001US-0313390P.
PR 28-AUG-2001; 2001US-0315470P.
PR 31-AUG-2001; 2001US-0316447P.
PR 07-SEP-2001; 2001US-0318115P.
PR 07-SEP-2001; 2001US-0318118P.
PR 12-SEP-2001; 2001US-0318740P.
PR 19-SEP-2001; 2001US-0323379P.
PR 18-OCT-2001; 2001US-0330245P.
PR 18-OCT-2001; 2001US-0330308P.
PR 14-NOV-2001; 2001US-0332701P.
XX
XX (CURA-) CURAGEN CORP.
XX

PI Tchernev VT, Spytek KA, Zerhusen BD, Patturajan M, Shimkets RA;
PI Li L, Gangolli EA, Padigaru M, Anderson DM, Rastelli L, Miller CE;
PI Gerlach VL, Taupier RJ, Gusev VY, Colman SD, Wolenc AR, Pena CE;
PI Furtak K, Grosse WM, Alsobrook JP, Lepley DM, Rieger DK, Burgess CE;
XX WPI; 2002-706998/76.
XX
XX New NOVX polypeptides and nucleic acids, useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
PT pharmacogenomics.
XX
XX Disclosure; SEQ ID NO 582; 1498pp; English.
XX
XX This invention relates to a novel nucleic acids, and encoded polypeptides
CC thereof, which have properties related to the stimulation of biochemical
CC or physiological responses in a cell, tissue, organ or organism.
CC Specifically, it refers to the use of biologically active fragments for
CC diagnostic and prognostic assays and furthermore in the treatment of
CC diverse pathological conditions. The present invention describes novel
CC human and murine NOVX proteins, as well as methods to modulate their
CC expression using antisense oligos, ribozymes and peptide nucleic acids.
CC The NOVX polypeptides, polynucleotides and antibodies are useful in
CC treating or preventing NOVX-associated disorders, e.g. cardiomyopathy,
CC atherosclerosis, cancer and diabetes. Furthermore, they may be used in
CC treating or preventing diseases such as inflammation, autoimmune
CC disorders, allergies, blood disorders, acquired immunodeficiency syndrome
CC (AIDS), obesity, asthma, immunoglobulin (Ig) nephropathy, cirrhosis,
CC arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy
CC and epilepsy. Accordingly, these molecules have many activities including
CC cyostatic, cardiant, antiinflammatory, immunosuppressive, antiallergic,
CC haemostatic, anti-HIV, antidiabetic, antiarteriosclerotic, anorectic,
CC antiasthmatic, nephrotropic, antiarthritic, hepatotropic,
CC neuroprotective, nootropic, antibacterial, virucide, antiparasitic,
CC relaxant and anticonvulsant. In addition, they are useful in screening
CC assays to identify small molecules that modulate or inhibit, for example,
CC neurogenesis, wound healing and angiogenesis. The nucleic acids are also
CC used as in chromosome mapping, tissue typing, preventive medicine and
CC pharmacogenomics. This polypeptide is a homologue of a human NOVX protein
CC of the invention.
XX
XX Sequence 304 AA;

Query Match 26.2%; Score 401; DB 5; Length 304;
Best Local Similarity 31.0%; Pred. No. 5.9e-33;
Matches 93; Conservative 62; Mismatches 125; Indels 20; Gaps 5;
QY 1 MAESEATPSTGPEFE-SKYFEFHGVRVLPFCRGKWEIANTFPVPSDVWIVYPKSGTS 59
Db 1 MAKIEKNAPTWEKKPELFNIMEVDGVTLLSKWEKVCNPFQAKPDDLILATYPSGTT 60
QY 60 LQEVYVLVSQADPDEIGLNMIDQLPVLE--YP---QPGLDIIKELTSPRLIKSHLPY 114
Db 61 WNHIELDMLNDGDVEKCKCAQTLDRHAFLELKFPHKEKPDLEFVLEWSSPOLIKTHLPS 120
QY 115 RFLPSDLNNGSKVITYMARNKDLVSVYQFHSRLTWSYRGTFQFCRRFNMNKLGVGS 174
Db 121 HLIPPSIWKENCKIVYVARNPKDCLVSYHFHRMASFMPDPQNLBEFYEFKMSGVVGS 180
QY 175 WFEHVQEFWEHRMDSNVLFLKYEDMHRDLVTWVEQLARFLGVSCDKAQLEALTEHQV 234
Db 181 WFDHVKGWAAKMDHRIILFYEDLKKPKKHEIHKVLEFKTWGSDVINKIVVHTSPDV 240
QY 235 ---DOCCNAEALP-----VGRGRVGLWKDIFTVSMNEKFDLVYKQWKGKCDLTF 280
Db 241 MKDNPMANHTAVPAHFNHSISKFKRKGMPGDKKHNFTVALNENFDKHYEKKAGSTLNF 300
RESULT 70
ABU12099
XX ABU12099 standard; protein; 305 AA.
ID
AC ABU12099;

XX 17-FEB-2003 (first entry)
XX Novel human sulfotransferase-like protein.
XX
XX Gamma-aminobutyric acid receptor-like protein; depression; stroke;
KW GABA receptor-like protein; Parkinson's disease; Huntington's disease;
KW Tourette's syndrome; amyotrophic lateral sclerosis; head trauma;
KW Alzheimer's disease; alcoholism; vigilance; anxiety; muscle tension;
KW epileptogenic activity; memory; cardiomyopathy; cancer; angiogenesis;
KW arrhythmogenic right ventricular dysplasia; renal disease; diabetes;
KW Epidermal growth factor like protein; leukaemia; lupus; anaemia; ulcer;
KW haematopoietic stem and progenitor cell like protein; cirrhosis;
KW sulfotransferase-like protein; cholangitis; hepatitis; hyperthyroidism;
KW developmental disorder; Syntaxin-like protein; myxoid liposarcoma;
KW asthma; Lambert-Eaton myasthenic syndrome; acute myeloidleukaemia;
KW transgenic animal.
XX
XX Homo sapiens.
XX
XX US2002123612-A1.
XX
XX 05-SEP-2002.
XX
XX 03-JUL-2001; 2001US-00898570.
XX
XX 19-APR-2000; 2000US-0198293P.
XX 20-APR-2000; 2000US-0198645P.
XX 25-APR-2000; 2000US-0199476P.
XX 26-APR-2000; 2000US-0199880P.
XX 26-APR-2000; 2000US-0200024P.
XX 26-APR-2000; 2000US-0200025P.
XX 09-JUN-2000; 2000US-0210809P.
XX 03-JUL-2000; 2000US-0215855P.
XX 17-JUL-2000; 2000US-0218591P.
XX 11-AUG-2000; 2000US-0224610P.
XX 27-FEB-2001; 2001US-0271814P.
XX
XX (GERL/) GERLACH V L.
XX (ELLE/) ELLERMAN K.
XX (MACD/) MACDOUGALL J R.
XX (SMT/) SMITHSON G.
XX
XX Gerlach VL, Ellerman K, Macdougall JR, Smithson G;
PI WPI; 2003-066815/06.
XX N-PSDB; ABX56478.
XX
XX Novel polypeptides and nucleic acids which are members of epidermal
PT growth factor, complement receptor families for diagnosing and treating
PT psychiatric conditions, depression, stroke, Alzheimer's and Parkinson's
PT disease.
XX
XX Claim 1; Page 35; 91pp; English.
XX
XX The invention describes an isolated POLYX (POLY1-17) polypeptide and its
CC variant. POLYX polypeptides (especially POLY5, POLY6 and POLY7), the
CC polynucleotides encoding them (I) and an anti-POLYX-antibody (III) are
CC useful for treating or preventing a pathology associated with POLYX
CC polypeptide in humans and for treating a syndrome associated with human
CC disease. POLYX polypeptide is also useful for identifying an agent that
CC binds to POLYX and a cell expressing POLYX is useful for identifying a
CC therapeutic agent for use in treatment of a pathology related to aberrant
CC expression or physiological interactions of the polypeptide. (III) is
CC useful for treating a pathological state in a mammal and for determining
CC the presence or amount of POLYX in a sample. POLY1-4 (GABA receptor-like
CC proteins) are useful for the treatment of psychiatric and medical
CC conditions, depression, stroke, Parkinson's disease, Huntington's
CC disease, Tourette's syndrome, amyotrophic lateral sclerosis, head trauma,
CC Alzheimer's disease, alcoholism, vigilance, anxiety, muscle tension,
CC epileptogenic activity and memory functions, cardiomyopathy and
CC arrhythmogenic right ventricular dysplasia. POLY5-8 (Epidermal growth
CC factor like proteins) may be useful for treating cancer, aberrant

CC angiogenesis, renal disease and diabetes. POLY12 (haematopoietic stem and
 CC progenitor cell like protein) may be useful for treatment of leukaemia,
 CC lupus and anaemia. POLY13 (sulfotransferase-like protein) may be useful
 CC for treating cirrhosis, cholangitis, hepatitis, ulcers, hyperthyroidism
 CC and developmental disorders. POLY14-16 (Syntaxin-like proteins) may be
 CC useful in treatment of Lambert-Eaton myasthenic syndrome, asthma, myxoid
 CC liposarcoma and acute myeloid leukaemia, and POLY 18 may be useful in
 CC treatment of cancers. Cells comprising (I) are useful for producing non-
 CC human transgenic animals which are useful for studying the function
 CC and/or activity of POLYX protein and for identifying and/or evaluating
 CC modulators of POLYX protein activity. This is the amino acid sequence of
 CC a novel human protein

XX Sequence 305 AA;

Query Match 26.1%; Score 399.5; DB 6; Length 305;
 Best Local Similarity 30.9%; Pred. No. 8.5e-33;
 Matches 93; Conservative 61; Mismatches 126; Indels 21; Gaps 5;
 QY 1 MAESEAETPTPGFE-SKYFEFGVRLPPFCRCGMKEIANFPVRPSDVWIVTPKSGTS 59
 Db 1 MAKTEKNAPTMEKKPELFNIMEVDGVTLLISKWEKVCNFQAKPDDLLILATYPKSGTT 60
 QY 60 LLOEVVYLVSGADPDEIGLNMIDQLPVLE--YP---QPGLDIKELTSPRLIKSHLPY 114
 Db 61 WWHIELDMLNDGVEKCKRAQTLDRHAFLEKLPKHEKPDLEFVLEMSSPQLIKTHLPS 120
 QY 115 RFLPSDLHGDSKVIYMARPKDLVSVYQFHRSLRTMSYRGTFQFCRRFMNDKLGYS 174
 Db 121 HLIPSIWKENCKIVYVARNPKDCLVSYVYHFRMASFMPDPQNLEEFYEKFMKSGKGF 180
 QY 175 WFEHVQFWEHRMDSNVLFKYEDEMHRDLVTMVEQLARFLGVSCDKAQLAEALTECHQLV 234
 Db 181 WFDHVKGWMAAKDHRILYLFYEDIKQNPKEIHKVLEFLEKTVSGDVINKIVHHTSPDV 240
 QY 235 ---DOCCNAEALPV-----GGRVGLWKDITVSNKFKDLVYKQKMGKCDLT 279
 Db 241 MKDNPMANHTAVPAHIFNHSISKFMKCGMPGDWKNHFTVALNENFDKHEKKGAGSTLN 300
 QY 280 F 280
 Db 301 F 301

RESULT 71

ID AAE05178

XX AAE05178 standard; protein; 304 AA.

AC AAE05178;

XX

DT 12-SEP-2001 (first entry)

XX Human drug metabolising enzyme (DME-9) protein.

XX Human; drug metabolising enzyme; DME-9; immunosuppressive; gene therapy;
 KW cystostatic; autoimmune disorder; inflammatory disorder; atherosclerosis;
 KW osteoporosis; eye disorder; hepatic tumour; Addison's disease; cretinism;
 KW rheumatoid arthritis; acquired immune deficiency syndrome; AIDS; anaemia;
 KW developmental disorder; endocrine disorder; iritis; acromegaly; epilepsy;
 KW thyrotoxicosis; pancreatic disorder; diabetes mellitus; obesity; adenoma;
 KW gastrointestinal disorder; nodular hyperplasia; conjunctivitis; glaucoma;
 KW acinic keratosis; metabolic disorder; dysphagia; anorexia; carcinoma;
 KW cell proliferative disorder.

XX Homo sapiens.

OS Key Location/Qualifiers

PH Binding-site 20..304

FT /label= PAPS_cofactor_binding_site.

XX WO200151638-A2.

XX 19-JUL-2001.

PD

XX 12-JAN-2001; 2001WO-US001174.
 PF
 XX 14-JAN-2000; 2000US-0176139P.
 PR 21-JAN-2000; 2000US-0177443P.
 PR 28-JAN-2000; 2000US-0178574P.
 XX (INCY-) INCYTE GENOMICS INC.
 XX Yang J, Baughn MR, Burford N, Au-Young J, Lu DAM, Reddy R;
 PI Ring HZ, Hillman JL, Yue H, Azimzai Y, Yao MG, Gandhi AR;
 PI Nguyen DB, Tang YT, Lal P, Bandman O;
 XX WPI; 2001-425874/45.
 DR N-PSDB; AAD09944.
 XX
 XX Drug metabolizing enzymes and encoding polynucleotides, useful for
 PT diagnosing, treating and/or preventing autoimmune, inflammatory, cell
 PT proliferative, developmental, endocrine, eye, metabolic, and
 PT gastrointestinal disorders.
 XX
 PS Claim 1; Page 143-144; 133pp; English.

XX The present sequence is human drug metabolising enzyme (DME-9) protein.
 CC Human DME and its nucleic acid molecule are useful for the diagnosis,
 CC treatment and prevention of disorders associated with increased or
 CC decreased expression of DME. Examples of such disorders include,
 CC autoimmune/inflammatory disorder such as acquired immune deficiency
 CC syndrome (AIDS), rheumatoid arthritis, osteoporosis; cell proliferative
 CC disorder such as actinic keratosis, atherosclerosis; developmental
 CC disorder such as epilepsy, anaemia; endocrine disorder such as
 CC acromegaly, cretinism, thyrotoxicosis; pancreatic disorder such as
 CC diabetes mellitus; eye disorder such as conjunctivitis, glaucoma, iritis;
 CC metabolic disorder such as Addison's disease, obesity; gastrointestinal
 CC disorder such as anorexia, dysphagia and hepatic tumours including
 CC nodular hyperplasia, adenomas and carcinomas. DME DNA is useful for
 CC creating 'knockin' humanised animals (pigs) or transgenic animals (mice
 CC or rats) to model human disease. DME DNA is also in useful is gene
 CC therapy. DME and its immunogenic fragments are useful for screening
 CC libraries of compounds in several drug screening assays

XX Sequence 304 AA;

Query Match 26.1%; Score 399; DB 4; Length 304;

Best Local Similarity 31.0%; Pred. No. 9.5e-33;

Matches 93; Conservative 61; Mismatches 126; Indels 20; Gaps 5;

QY 1 MAESEAETPTPGFE-SKYFEFGVRLPPFCRCGMKEIANFPVRPSDVWIVTPKSGTS 59
 Db 1 MAKTEKNAPTMEKKPELFNIMEVDGVTLLISKWEKVCNFQAKPDDLLILATYPKSGTT 60
 QY 60 LLOEVVYLVSGADPDEIGLNMIDQLPVLE--YP---QPGLDIKELTSPRLIKSHLPY 114
 Db 61 WWHIELDMLNDGVEKCKRAQTLDRHAFLEKLPKHEKPDLEFVLEMSSPQLIKTHLPS 120
 QY 115 RFLPSDLHGDSKVIYMARPKDLVSVYQFHRSLRTMSYRGTFQFCRRFMNDKLGYS 174
 Db 121 HLIPSIWKENCKIVYVARNPKDCLVSYVYHFRMASFMPDPQNLEEFYEKFMKSGKVV 180
 QY 175 WFEHVQFWEHRMDSNVLFKYEDEMHRDLVTMVEQLARFLGVSCDKAQLAEALTECHQLV 234
 Db 181 WFDHVKGWMAAKDHRILYLFYEDIKQNPKEIHKVLEFLEKTVSGDVINKIVHHTSPDV 240
 QY 235 ---DOCCNAEALPV-----VGRVGLWKDITVSNKFKDLVYKQKMGKCDLT 280
 Db 241 MKDNPMANHTAVPAHIFNHSISKFMKCGMPGDWKNHFTVALNENFDKHEKKGAGSTLN 300
 RESULT 72
 ID ADD12583
 XX ADD12583 standard; protein; 261 AA.
 XX AC ADD12583;

XX	DT	01-JAN-2004 (first entry)	
XX	DE	Human ENZM-43 protein SEQ ID NO:43.	
XX	KW	human; enzyme; ENZM; cytosolic; antiarteriosclerotic; anti-HIV; antiatherogenic; cerebroprotective; antiparkinsonian; anticonvulsant; neurotropic; neuroprotective; antiinflammatory; ophthalmological; antithyroid; antiarthritic; antibacterial; virucide; protozoacide; antiparasitic; fungicide; anorectic; cardiant; hypotensive; antinfertility; hepatotropic; gene therapy; autoimmune disorder; inflammatory disorder; AIDS; allergy; atopic dermatitis; arthritis; thyroiditis; infection; metabolic disorder; obesity; reproductive disorder; infertility; neurological disorder; Parkinson's disease; Alzheimer's disease; stroke; epilepsy; cardiovascular disorder; myocardial infarction; hypertension; eye disorder; cell proliferative disease; cancer; atherosclerosis; hepatitis.	
OS	KW	Homo sapiens.	
XX	PN	WO2003072729-A2.	
XX	PD	04-SEP-2003.	
XX	PF	21-FEB-2003; 2003WO-US005478.	
XX	PR	22-FEB-2002; 2002US-0359513P.	
XX	PR	19-MAR-2002; 2002US-0365795P.	
XX	PA	(INCY-) INCYTE GENOMICS INC.	
PI	PI	Sprague WW, Kable AE, Lee S, Griffin JA, Becha SD, Lee EA; Hafalia AJA, Khare R, Emerling BM, Marquis JP, Rankumar J; Elliott VS, Richardson TW, Baughn MR, Jin P, Chien D, Hawkins PR; Wilton AD, Chawla NK, Tran UK, Lee SY, Zeburjadian Y, Jiang X; Jackson AA, Bhatia UG, Burrill JD, Blake JJ, Ho A, Zheng W; WPI; 2003-731608/69.	
DR	N-PSDB	ADD12640.	
XX	PT	New human enzymes and polynucleotides, useful for diagnosing, preventing or treating diseases or conditions associated with aberrant enzyme expression, e.g. cancer, hepatitis, AIDS, atherosclerosis, obesity, infection or stroke.	
PS	PS	Claim 1; SEQ ID NO 43; 426pp; English.	
CC	CC	The present sequence represents a human enzyme designated ENZM-43. Human enzymes of the present invention have cytosolic, antiarteriosclerotic, anti-HIV, antiatherogenic, cerebroprotective, antiparkinsonian, anticonvulsant, neurotropic, neuroprotective, antiinflammatory, ophthalmological, antithyroid, antiarthritic, antibacterial, virucide, protozoacide, antiparasitic, fungicide, anorectic, cardiant, hypotensive, antinfertility and hepatotropic activities, and can be used in gene therapy. The human enzymes and polynucleotides encoding them can be used in diagnosing, preventing or treating diseases or conditions associated with the decreased expression or overexpression of the enzymes, such as autoimmune/inflammatory disorders (AIDS, allergies, atopic dermatitis, arthritis or thyroiditis), infections (e.g. bacterial, viral, parasitic, protozoal or fungal), metabolic disorders (e.g. obesity), reproductive disorders (e.g. infertility), neurological disorders (Parkinson's disease, Alzheimer's disease, stroke, epilepsy), cardiovascular disorders (e.g. myocardial infarction or hypertension), eye disorders, or cell proliferative diseases (e.g. cancer, atherosclerosis, hepatitis). They are also useful in assessing the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of the enzymes.	
XX	XX	Sequence 261 AA;	
XX	XX	Query Match 26.0%; Score 397.5; DB 7; Length 261;	
XX	XX	Best Local Similarity 30.4%; Pred. No. 1.1e-32;	
XX	XX	Matches 85; Conservative 53; Mismatches 91; Indels 51; Gaps 3;	

QY	18	KYFEHGVRLPFCRGKMEIEANFVRPSDVWIVTYPKSGTSLLOEVVYLVSQGADPDEI	77
Db	12	KLKEVEGTLLOPATVDNMSQIQSFEAKPDDLICTYPRAGTTWIOEIV	59
QY	78	GLMNIDEQLPVLEYQPGLDITIKELTSPRLIKSHLPYRFLPSDLHNGSKVLYMARNPKD	137
Db	60	-----EKAKAMPSPRILKTHLSTQLLPSPFWENCKFLYVARNAKD	100
QY	138	LVSYYQPHRSRLRTMSYRGTFQFECRRFMNDKLGYSWFHVEHFRMDSNVLPFLKYE	197
Db	101	CMVSYHFEQRMNHLDPDGTWEEYFETFGKVGWGSFDFHVKGWEMKDRHQLFLFYE	160
QY	198	DHMRDLVTWEOQLARFLGVSCDKAQLEALTEH	240
Db	161	DIKRPDKEIRKVMQFMGKKVDYTLDKIVQTSFEKMKENPMTNRTSVKSKILDQSISS	220
QY	241	EALPVGRGRVGLWKDIFTVSMNEKEDLVYKOKMGKCDLTF	280
Db	221	F---MRKGTVDGDKNHFVAQNERFDEYRRKMGTSINF	257
RESULT	73		
ADII16670			
ID	ADII16670	standard; protein; 295 AA.	
XX	AC	ADII16670;	
XX	DT	15-APR-2004 (first entry)	
XX	DE	Human NOVX protein to treat human pathological conditions SeqID206.	
KW	KW	human; NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes; inflammation; autoimmune disorder; allergy; blood disorder; acquired immunodeficiency syndrome; AIDS; obesity; asthma; immunoglobulin (Ig)A nephropathy; cirrhosis; arthritis; Alzheimer's disease; infection; stroke; muscular dystrophy; epilepsy; cytostatic; cardiant; antiinflammatory; immunosuppressive; antiallergic; haemostatic; anti-HIV; antidiabetic; antiarteriosclerotic; anorectic; antiasthmatic; nephrotropic; antiarthritic; hepatotropic; neuroprotective; neurotropic; antibacterial; virucide; antiparasitic; relaxant; anticonvulsant; neurogenesis; wound healing; angiogenesis; chromosome mapping; tissue typing; pharmacogenomic.	
OS	OS	Homo sapiens.	
XX	PN	WO200268649-A2.	
XX	PD	06-SEP-2002.	
XX	PF	31-JAN-2002; 2002WO-US002785.	
XX	PR	31-JAN-2001; 2001US-0265395P.	
PR	PR	31-JAN-2001; 2001US-0265412P.	
PR	PR	31-JAN-2001; 2001US-0265514P.	
PR	PR	31-JAN-2001; 2001US-0265517P.	
PR	PR	02-FEB-2001; 2001US-0266406P.	
PR	PR	05-FEB-2001; 2001US-0266767P.	
PR	PR	07-FEB-2001; 2001US-0266975P.	
PR	PR	07-FEB-2001; 2001US-0267057P.	
PR	PR	08-FEB-2001; 2001US-0267459P.	
PR	PR	09-FEB-2001; 2001US-0267823P.	
PR	PR	15-FEB-2001; 2001US-0268974P.	
PR	PR	26-FEB-2001; 2001US-0271664P.	
PR	PR	27-FEB-2001; 2001US-0271839P.	
PR	PR	27-FEB-2001; 2001US-0271855P.	
PR	PR	02-MAR-2001; 2001US-0272788P.	
PR	PR	02-MAR-2001; 2001US-0273046P.	
PR	PR	14-MAR-2001; 2001US-0275925P.	
PR	PR	14-MAR-2001; 2001US-0275947P.	
PR	PR	14-MAR-2001; 2001US-0275950P.	
PR	PR	14-MAR-2001; 2001US-0275989P.	
PR	PR	15-MAR-2001; 2001US-0276448P.	


```
XX PR 12-SEP-2002; 2002US-0410259P.
XX PR 12-SEP-2002; 2002US-0410260P.
XX PA (INCY-) INCYTE CORP.
XX PI Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
XX PI Harthshorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
XX PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;
XX PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
XX PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
XX PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;
XX PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
XX PI Patury S, Shi X, Suarez CJ;
XX PF WPI; 2004-329368/30.
XX DR N-PSDB; ACN42660.
XX PT New diagnostic and therapeutic polynucleotides and polypeptides, useful
XX PT in diagnosing a condition, disease or disorder associated with human
XX PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
XX PT in gene mapping.
XX PS Claim 27; Page; 190pp; English.
XX CC The invention relates to novel diagnostic and therapeutic polynucleotides
XX CC selected from one of the 2722 sequences defined in the specification. A
XX CC polynucleotide of the invention may have a use in gene therapy. The human
XX CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
XX CC used to diagnose a particular condition, disease or disorder associated
XX CC with human molecules, e.g. cell proliferative disorders,
XX CC autoimmune/inflammatory disorder, developmental disorder, endocrine
XX CC disorder, neurological disorders, gastrointestinal disorders, or
XX CC infections caused by virus, bacteria, fungi or parasite. The dithp
XX CC molecules may also be used in genetic mapping, in identifying individuals
XX CC from minute biological samples, in detecting single nucleotide
XX CC polymorphisms, as molecular weight markers, and for somatic or germline
XX CC gene therapy. The present sequence represents a dithp protein of the
XX CC invention. Note: The sequence data for this patent is not represented in
XX CC the printed specification, but was obtained in electronic format directly
XX CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX SQ Sequence 262 AA;
XX Query Match 25.1%; Score 384.5; DB 8; Length 262;
XX Best Local Similarity 30.5%; Pred. No. 2.6e-31;
XX Matches 84; Conservative 49; Mismatches 91; Indels 51; Gaps 4;
XX QY 24 GVRLPFPCRGKMEETANFVRPSDVWIVTYPKSGTSLLOEVVYLVSQGADPEIGLMNID 83
XX DB 17 GVPLIKYPAEALGPLQSQFARPDLLINTYPKSGTTWVSQILDMIYQGGDLKCNRAPIY 76
XX QY 84 EQLPVLEYPPQGLDIIKELTSPLRIKSHLPYRFLPSDLHNGDSKVIYMARNPKDLVSVY 143
XX DB 77 VRVPLEVNDP-----GEPGSGVYVARNPKDVAVSY 107
XX QY 144 QFHRSLRTMSYRGTFQEFRCRPMNDKLGYSWFEHVBQFWEHRMDSNVLFKLYEDMHRDL 203
XX DB 108 HFHRMEKAPPEPTGTFDSLEKFMAGEVSGYSQYHVEWELSRTHPYLYLFYEDMKNP 167
XX QY 204 VTMVQLARFLGVSCDKAQLEALTEHC-----HQLVDQCNAEALP-V 245
XX DB 168 KREIKILFVGRSGSLPEETMDPMVQHTSFKEKKMKNPMNTYTVPQELMDHSIS-----PPM 223
XX QY 246 GRGRVGLMKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
XX DB 224 RKGWAGDWKTTFTVAQNERFDADYAEKNAGCSLSF 258
XX RESULT 79
XX ABM82616
XX ID ABM82616 standard; protein; 262 AA.
```

```
AC ABM82616;
XX 18-NOV-2004 (first entry)
XX Human diagnostic and therapeutic pprotein SEQ ID NO:2865.
XX gene therapy; human diagnostic and therapeutic polynucleotide; dithp.
XX Homo sapiens.
XX WO2004023973-A2.
XX 25-MAR-2004.
XX 12-SEP-2003; 2003WO-US028227.
XX 12-SEP-2002; 2002US-0410259P.
XX 12-SEP-2002; 2002US-0410260P.
XX (INCY-) INCYTE CORP.
XX Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
XX Harthshorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
XX Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;
XX Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
XX Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
XX Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;
XX Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
XX Patury S, Shi X, Suarez CJ;
XX WPI; 2004-329368/30.
XX N-PSDB; ACN41268.
XX New diagnostic and therapeutic polynucleotides and polypeptides, useful
XX in diagnosing a condition, disease or disorder associated with human
XX molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
XX in gene mapping.
XX Claim 27; Page; 190pp; English.
XX The invention relates to novel diagnostic and therapeutic polynucleotides
XX selected from one of the 2722 sequences defined in the specification. A
XX polynucleotide of the invention may have a use in gene therapy. The human
XX diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
XX used to diagnose a particular condition, disease or disorder associated
XX with human molecules, e.g. cell proliferative disorders,
XX autoimmune/inflammatory disorder, developmental disorder, endocrine
XX disorder, neurological disorders, gastrointestinal disorders, or
XX infections caused by virus, bacteria, fungi or parasite. The dithp
XX molecules may also be used in genetic mapping, in identifying individuals
XX from minute biological samples, in detecting single nucleotide
XX polymorphisms, as molecular weight markers, and for somatic or germline
XX gene therapy. The present sequence represents a dithp protein of the
XX invention. Note: The sequence data for this patent is not represented in
XX the printed specification, but was obtained in electronic format directly
XX from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX SQ Sequence 262 AA;
XX Query Match 25.1%; Score 384.5; DB 8; Length 262;
XX Best Local Similarity 30.5%; Pred. No. 2.6e-31;
XX Matches 84; Conservative 49; Mismatches 91; Indels 51; Gaps 4;
XX QY 24 GVRLPFPCRGKMEETANFVRPSDVWIVTYPKSGTSLLOEVVYLVSQGADPEIGLMNID 83
XX DB 17 GVPLIKYPAEALGPLQSQFARPDLLINTYPKSGTTWVSQILDMIYQGGDLKCNRAPIY 76
XX QY 84 EQLPVLEYPPQGLDIIKELTSPLRIKSHLPYRFLPSDLHNGDSKVIYMARNPKDLVSVY 143
XX DB 77 VRVPLEVNDP-----GEPGSGVYVARNPKDVAVSY 107
XX QY 144 QFHRSLRTMSYRGTFQEFRCRPMNDKLGYSWFEHVBQFWEHRMDSNVLFKLYEDMHRDL 203
```

Db 108 HPHRMEKAHPGPTWDSFLEKFMAGEVSGSWYQHVOEWELSRTHPVLYLFYEDMKENP 167

Qy 204 VTMVQLARFLGVSCDKAQLEALTEHC-----HQLVDQCCNAEALP-V 245

Db 168 KREIQKILEFVGRSLPEETMDFMVQHTSPKMKKNPMNTYTTVPOELMDHSIS-----PFM 223

Qy 246 GGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280

Db 224 RKGWAGDWKTTTVAQNERFDADYAEKMGAGCSLSF 258

RESULT 80

ABM84005

ID ABM84005 standard; protein; 262 AA.

AC ABM84005;

XX

XX 18-NOV-2004 (first entry)

XX Human diagnostic and therapeutic pprotein SEQ ID NO:4254.

XX gene therapy; human diagnostic and therapeutic polynucleotide; dithp.

XX Homo sapiens.

XX WO2004023973-A2.

XX 25-MAR-2004.

XX 12-SEP-2003; 2003WO-US028227.

XX 12-SEP-2002; 2002US-0410259P.

PR 12-SEP-2002; 2002US-0410260P.

XX (INCY-) INCYTE CORP.

XX Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;

PI Harthshorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;

PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;

PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;

PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;

PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;

PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;

PI Patury S, Shi X, Suarez CJ;

XX WPI; 2004-329368/30.

DR N-PSDB; ACN42657.

XX New diagnostic and therapeutic polynucleotides and polypeptides, useful

PT in diagnosing a condition, disease or disorder associated with human

PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or

PT in gene mapping.

XX Claim 27; Page; 190pp; English.

XX The invention relates to novel diagnostic and therapeutic polynucleotides

CC selected from one of the 2722 sequences defined in the specification. A

CC polynucleotide of the invention may have a use in gene therapy. The human

CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be

CC used to diagnose a particular condition, disease or disorder associated

CC with human molecules, e.g. cell proliferative disorders,

CC autoimmune/inflammatory disorder, developmental disorder, endocrine

CC disorder, neurological disorders, gastrointestinal disorders, or

CC infections caused by virus, bacteria, fungi or parasite. The dithp

CC molecules may also be used in genetic mapping, in identifying individuals

CC from minute biological samples, in detecting single nucleotide

CC polymorphisms, as molecular weight markers, and for somatic or germline

CC gene therapy. The present sequence represents a dithp protein of the

CC invention. Note: The sequence data for this patent is not represented in

CC the printed specification, but was obtained in electronic format directly

CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm

XX Sequence 262 AA;

Query Match 25.1%; Score 384.5; DB 8; Length 262;

Best Local Similarity 30.5%; Pred. No. 2.6e-31;

Matches 84; Conservative 49; Mismatches 91; Indels 51; Gaps 4;

Qy 24 GVRLPFFCRGKMEETANFPVRPSDVWIVTYPKSGTSLLOEVVYLVSOQADPDEIGLMNID 83

Db 17 GVPLIKYFAEALGPLQSFQARPDLLINTYPKSGTSLLOEVVYLVSOQADPDEIGLMNID 76

Qy 84 EQLPVLEYPOQGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYNARPNKDLVVSYY 143

Db 77 VRVPFLEVNDP-----GEPSGVYVARNPKOVAVSYY 107

Qy 144 QHRSRLRTMSYRGTFQBFRCRPFMDKLGYSWFEHVFQFWEHRMDSNVLFKLYEDMHRDL 203

Db 108 HPHRMEKAHPGPTWDSFLEKFMAGEVSGSWYQHVOEWELSRTHPVLYLFYEDMKENP 167

Qy 204 VTMVQLARFLGVSCDKAQLEALTEHC-----HQLVDQCCNAEALP-V 245

Db 168 KREIQKILEFVGRSLPEETMDFMVQHTSPKMKKNPMNTYTTVPOELMDHSIS-----PFM 223

Qy 246 GGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280

Db 224 RKGWAGDWKTTTVAQNERFDADYAEKMGAGCSLSF 258

RESULT 81

ABM82696

ID ABM82696 standard; protein; 253 AA.

XX AC ABM82696;

XX 18-NOV-2004 (first entry)

XX Human diagnostic and therapeutic pprotein SEQ ID NO:2945.

XX gene therapy; human diagnostic and therapeutic polynucleotide; dithp.

XX Homo sapiens.

XX WO2004023973-A2.

XX 25-MAR-2004.

XX 12-SEP-2003; 2003WO-US028227.

PR 12-SEP-2002; 2002US-0410259P.

PR 12-SEP-2002; 2002US-0410260P.

XX (INCY-) INCYTE CORP.

XX Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;

PI Harthshorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;

PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;

PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;

PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;

PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;

PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;

PI Patury S, Shi X, Suarez CJ;

XX WPI; 2004-329368/30.

DR N-PSDB; ACN41348.

XX New diagnostic and therapeutic polynucleotides and polypeptides, useful

PT in diagnosing a condition, disease or disorder associated with human

PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or

PT in gene mapping.

XX Claim 27; Page; 190pp; English.

XX The invention relates to novel diagnostic and therapeutic polynucleotides

CC selected from one of the 2722 sequences defined in the specification. A

CC polynucleotide of the invention may have a use in gene therapy. The human

CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be

CC used to diagnose a particular condition, disease or disorder associated

CC with human molecules, e.g. cell proliferative disorders,

CC autoimmune/inflammatory disorder, developmental disorder, endocrine

CC disorder, neurological disorders, gastrointestinal disorders, or

CC infections caused by virus, bacteria, fungi or parasite. The dithp

CC molecules may also be used in genetic mapping, in identifying individuals

CC from minute biological samples, in detecting single nucleotide

CC polymorphisms, as molecular weight markers, and for somatic or germline

CC gene therapy. The present sequence represents a dithp protein of the

CC invention. Note: The sequence data for this patent is not represented in

CC the printed specification, but was obtained in electronic format directly

CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm

XX Sequence 262 AA;

PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241222P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.

PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246603P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 17-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0251990P.
PR 05-JAN-2001; 2001US-0254097P.
XX 05-JAN-2001; 2001US-0259678P.
PA (HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI, 2001-488787/53.
N-PSDB; AAS30231.

New polynucleotides and polypeptides, useful for diagnosing, treating, preventing or prognosing e.g. kidney, cardiovascular, blood, electrolyte imbalance or neoplastic disorders, autoimmune diseases, cancers.

Claim 1; SEQ ID NO 149; 506pp; English.

The invention relates to novel nucleic acids and polypeptides useful for diagnosing, treating, preventing and/or prognosing disorders related to these polypeptides. The polynucleotides are especially useful in the diagnosis, prognosis, prevention and/or treatment of diseases which include kidney disorders (e.g. renal failure or nephritis), cardiovascular disorders (e.g. hypertension or myocardial infarction), blood disorders (e.g. anaemia or blood coagulation disorders), electrolyte imbalance disorders (e.g. hyponatraemia or hyperkalaemia), neoplastic disorders (e.g. nephroma or renal cell cancer), autoimmune diseases, cancers, inflammatory diseases, reproductive system disorders, endocrine disorders, neural activity and neurological disorders, wound

CC healing and respiratory disorders. AAU18644-AAU18715 represent the novel
CC human renal and cardiovascular-associated amino acid sequences of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at: ftp.wipo.int/pub/published_pct_sequences

XX Sequence 76 AA;

Query Match 25.0%; Score 382; DB 4; Length 76;

Best Local Similarity 95.9%; Pred. No. 8.2e-32;

Matches 71; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 90 EYPQGLDIIKELTSPRIKSHLPYRFLPSDLHNGDSKVIYMARPKDLVSYQFHRSL 149
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Db 1 KYQPGLDIIKELTSPRIKSHLPYRFLPSDLHNGDSKVIYMARPKDLVSYQFHRSL 60

QY 150 RTMSYRGTFQBFQR 163

Db 61 RTMSYRGTFQBFQR 74

RESULT 85

ABBI14904

ID ABBI14904 standard; protein; 76 AA.

XX AC ABBI14904;

XX XX

XX 23-JAN-2002 (first entry)

XX DE Human nervous system related polypeptide SEQ ID NO 3561.

XX Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery;
KW antiparkinsonian; antisticking; antianaemic; antiarthritic; cancer;
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine.

XX OS Homo sapiens.

XX XX WO200159063-A2.

XX PD 16-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US001334.

XX PR 31-JAN-2000; 2000US-0179065P.

XX PR 04-FEB-2000; 2000US-0180628P.

XX PR 24-FEB-2000; 2000US-0184664P.

XX PR 02-MAR-2000; 2000US-0186350P.

XX PR 16-MAR-2000; 2000US-0189874P.

XX PR 17-MAR-2000; 2000US-0190076P.

XX PR 18-APR-2000; 2000US-019123P.

XX PR 19-MAY-2000; 2000US-0205515P.

XX PR 07-JUN-2000; 2000US-0209467P.

XX PR 28-JUN-2000; 2000US-0214886P.

XX PR 30-JUN-2000; 2000US-0215135P.

XX PR 07-JUL-2000; 2000US-0216647P.

XX PR 07-JUL-2000; 2000US-0216880P.

XX PR 11-JUL-2000; 2000US-0217487P.

XX PR 11-JUL-2000; 2000US-0217496P.

XX PR 14-JUL-2000; 2000US-0218290P.

XX PR 26-JUL-2000; 2000US-0220963P.

XX PR 26-JUL-2000; 2000US-0220964P.

XX PR 14-AUG-2000; 2000US-0224518P.

XX PR 14-AUG-2000; 2000US-0224519P.

XX PR 14-AUG-2000; 2000US-0225213P.

XX PR 14-AUG-2000; 2000US-0225214P.

XX PR 14-AUG-2000; 2000US-0225266P.

XX PR 14-AUG-2000; 2000US-0225267P.

XX PR 14-AUG-2000; 2000US-0225268P.

XX PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231245P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234224P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236377P.
PR 29-SEP-2000; 2000US-0236378P.
PR 29-SEP-2000; 2000US-0236386P.
PR 29-SEP-2000; 2000US-0236389P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 20-OCT-2000; 2000US-0242221P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.

PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0223343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0231418P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235835P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0235802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.

PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249298P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 PR 17-JAN-2001; 2001US-00764863.

XX (ROSE/) ROSEN C A.
 PA (RUBE/) RUBEN S M.
 PA (BARA/) BARASH S C.

XX PI Rosen CA, Ruben SM, Barash SC;
 XX WPI; 2003-416807/39.
 DR N-PSDB; ACD01486.

XX New nucleic acids and polypeptides, useful for diagnosing, prognosing,
 PT preventing and/or treating e.g. kidney, cardiovascular, blood,
 PT electrolyte imbalance, neoplastic, respiratory, or neurological diseases
 or disorders.

XX Claim 11; Page 304; 363pp; English.

XX The invention relates to isolated nucleic acids encoding novel
 CC polypeptides. The nucleic acids are useful for chromosome mapping, for
 CC radiation hybrid mapping, for detection of cancer, in gene therapy, for
 CC identifying individuals from minute biological samples, as an alternative
 CC to restriction fragment length polymorphism (RFLP) analysis, in forensic
 CC biology and as hybridisation probes for differential identification of
 CC tissues or cell types present in a biological sample. Compositions
 CC comprising the polynucleotides, polypeptides and antibodies specific for
 CC the polypeptides may be used in the diagnosis, prognosis, prevention
 CC and/or treatment of kidney disorders (e.g. glomerulonephritis, urinary
 CC tract infections, chronic nephritis), cardiovascular disorders (e.g.
 CC arrhythmias, carotid heart disease, endocarditis), blood disorders
 CC (e.g. thrombosis, anaemia, haemoglobin abnormalities), electrolyte
 CC imbalance, neoplastic disorders (e.g. cancers), respiratory disorders
 CC (e.g. acute rhinitis, sinusitis, pharyngitis) and neurological disorders
 CC (e.g. Alzheimer's disease, Parkinson's disease, Huntington's disease).
 CC Sequences ABU97259-ABU97330 represent human polypeptides of the invention

XX Sequence 76 AA;

Query Match 25.0%; Score 382; DB 6; Length 76;

Best Local Similarity 95.9%; Pred. No. 8.2e-32;

Matches 71; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 90 EYPQGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARPKDLVWSYQFHRSL 149

DB 1 KYQPGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARPKDLVWSYQFHRSL 60

QY	150	RTMSYRGTFQBFRCR	163	
DB	61	RTMSYRGXFQBFRCR	74	
RESULT 87				
ABM84007				
ID	ABM84007	standard; protein; 254 AA.		
XX	AC	ABM84007;		
XX	18-NOV-2004	(first entry)		
DT	18-NOV-2004	(first entry)		
DE	Human diagnostic and therapeutic pprotein	SEQ ID NO:4256.		
KW	gene therapy; human diagnostic and therapeutic polynucleotide; dithp.			
OS	Homo sapiens.			
XX	WO2004023973-A2.			
PD	25-MAR-2004.			
PF	12-SEP-2003; 2003WO-US028227.			
XX	12-SEP-2002; 2002US-0410259P.			
PR	12-SEP-2002; 2002US-0410260P.			
XX	(INCY-) INCYTE CORP.			
PI	Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;			
PI	Harthshorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;			
PI	Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;			
PI	Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;			
PI	Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;			
PI	Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;			
PI	Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;			
PI	Patury S, Shi X, Suarez CJ;			
XX	WPI; 2004-329368/30.			
DR	N-PSDB; ACN42659.			
XX	New diagnostic and therapeutic polynucleotides and polypeptides, useful			
PT	in diagnosing a condition, disease or disorder associated with human			
PT	molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or			
PT	in gene mapping.			
XX	Claim 27; Page: 190pp; English.			
XX	The invention relates to novel diagnostic and therapeutic polynucleotides			
CC	selected from one of the 2722 sequences defined in the specification. A			
CC	polynucleotide of the invention may have a use in gene therapy. The human			
CC	diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be			
CC	used to diagnose a particular condition, disease or disorder associated			
CC	with human molecules, e.g. cell proliferative disorders,			
CC	autoimmune/inflammatory disorder, developmental disorder, endocrine			
CC	disorder, neurological disorders, gastrointestinal disorders, or			
CC	infections caused by virus, bacteria, fungi or parasite. The dithp			
CC	molecules may also be used in genetic mapping, in identifying individuals			
CC	from minute biological samples, in detecting single nucleotide			
CC	polymorphisms, as molecular weight markers, and for somatic or germline			
CC	gene therapy. The present sequence represents a dithp protein of the			
CC	invention. Note: The sequence data for this patent is not represented in			
CC	the printed specification, but was obtained in electronic format directly			
CC	from WIPO at www.wipo.int/pct/en/sequences/listing.htm			
XX	Sequence 254 AA;			
SQ				
Query Match				
Best Local Similarity 24.6%; Score 376.5; DB 8; Length 254;				
Matches 83; Conservative 48; Mismatches 85; Indels 59; Gaps 4;				
QY	24	GVRLLPFCKRKEEIANFFPRSDVWIVTPKSGTSLQEVVLYVSQGADPDEIGLMNID	83	

CC molecules may also be used in genetic mapping, in identifying individuals
CC from minute biological samples, in detecting single nucleotide
CC polymorphisms, as molecular weight markers, and for somatic or germline
CC gene therapy. The present sequence represents a dithp protein of the
CC invention. Note: The sequence data for this patent is not represented in
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm

XX Sequence 254 AA;

Query Match 24.6%; Score 376.5; DB 8; Length 254;
Best Local Similarity 30.2%; Pred. No. 1.7e-30;
Matches 83; Conservative 48; Mismatches 85; Indels 59; Gaps 4;

QY 24 GVRLLPPFCRGKMEETANFVRPSDVWIVTPKSGTSLQEVVYLVSQADPDEIGLNMID 83
DB 17 GVPLIKYFAEALGPLQSFQARPDDLLINTYPKSGTTWVSQILDIMYQGGDLEKCNRAPIY 76
QY 84 EQLPVLEYPQGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYNARPNKDLVVSYY 143
DB 77 VRVPFLE-----VVYVARNPKDVAVSYY 99
QY 144 QFHESLRTMSYRGTFQBFRCRFMNDKLGYSWFHVBQFWEHRMDSNVLFKYEEDMHRDL 203
DB 100 HFHRMEKAHPGPTWDSFLEKFMAGEVSYGSWYQHVQEWELSRTHPVLVLYFYEDMKENP 159
QY 204 VTMVEQLARFLGVSCKAQLAETHC-----HQLVDQCCNAEALP-V 245
DB 160 KREIQKILEFVGRSLPPEETDMFVQHTSFKEMKKNPMTNYTTPQELMDHSIS-----PFM 215
QY 246 GRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
DB 216 RKG MAGDWKTTFTVAQNERFDADYAEKMGAGCSLSF 250

RESULT 89

ABM84006
ID ABM84006 standard; protein; 254 AA.
XX
AC ABM84006;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human diagnostic and therapeutic pprotein SEQ ID NO:4255.
XX
KW gene therapy; human diagnostic and therapeutic polynucleotide; dithp.
XX
OS Homo sapiens.
XX
PN WO2004023973-A2.
XX
PD 25-MAR-2004.
XX
PF 12-SEP-2003; 2003WO-US028227.
XX
PR 12-SEP-2002; 2002US-0410259P.
XX
PR 12-SEP-2002; 2002US-0410260P.
XX
PA (INCY-) INCYTE CORP.

PI Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
PI Harthshorne TA, Suchorolski WT, Altus CM, Birtts SJ, Elder LV;
PI Mooney EM, Deleane AM, Panesar IS, Banville SC, Reddy TP;
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirtton ES;
PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
PI Patury S, Shi X, Suarez CJ;
XX
XX WPI; 2004-329368/30.
DR N-PSDB; ACN42658.
XX
XX New diagnostic and therapeutic polynucleotides and polypeptides, useful

PT in diagnosing a condition, disease or disorder associated with human
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT in gene mapping.

XX Claim 27; Page: 190pp; English.

XX The invention relates to novel diagnostic and therapeutic polynucleotides
CC selected from one of the 2722 sequences defined in the specification. A
CC polynucleotide of the invention may have a use in gene therapy. The human
CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
CC used to diagnose a particular condition, disease or disorder associated
CC with human molecules, e.g. cell proliferative disorders,
CC autoimmune/inflammatory disorder, developmental disorder, endocrine
CC disorder, neurological disorders, gastrointestinal disorders, or
CC infections caused by virus, bacteria, fungi or parasite. The dithp
CC molecules may also be used in genetic mapping, in identifying individuals
CC from minute biological samples, in detecting single nucleotide
CC polymorphisms, as molecular weight markers, and for somatic or germline
CC gene therapy. The present sequence represents a dithp protein of the
CC invention. Note: The sequence data for this patent is not represented in
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm

XX Sequence 254 AA;

Query Match 24.6%; Score 376.5; DB 8; Length 254;
Best Local Similarity 30.2%; Pred. No. 1.7e-30;
Matches 83; Conservative 48; Mismatches 85; Indels 59; Gaps 4;

QY 24 GVRLLPPFCRGKMEETANFVRPSDVWIVTPKSGTSLQEVVYLVSQADPDEIGLNMID 83
DB 17 GVPLIKYFAEALGPLQSFQARPDDLLINTYPKSGTTWVSQILDIMYQGGDLEKCNRAPIY 76
QY 84 EQLPVLEYPQGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYNARPNKDLVVSYY 143
DB 77 VRVPFLE-----VVYVARNPKDVAVSYY 99
QY 144 QFHESLRTMSYRGTFQBFRCRFMNDKLGYSWFHVBQFWEHRMDSNVLFKYEEDMHRDL 203
DB 100 HFHRMEKAHPGPTWDSFLEKFMAGEVSYGSWYQHVQEWELSRTHPVLVLYFYEDMKENP 159
QY 204 VTMVEQLARFLGVSCKAQLAETHC-----HQLVDQCCNAEALP-V 245
DB 160 KREIQKILEFVGRSLPPEETDMFVQHTSFKEMKKNPMTNYTTPQELMDHSIS-----PFM 215
QY 246 GRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
DB 216 RKG MAGDWKTTFTVAQNERFDADYAEKMGAGCSLSF 250

RESULT 90

ABM82619
ID ABM82619 standard; protein; 254 AA.
XX
AC ABM82619;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human diagnostic and therapeutic pprotein SEQ ID NO:2868.
XX
KW gene therapy; human diagnostic and therapeutic polynucleotide; dithp.
XX
OS Homo sapiens.
XX
PN WO2004023973-A2.
XX
PD 25-MAR-2004.
XX
PF 12-SEP-2003; 2003WO-US028227.
XX
PR 12-SEP-2002; 2002US-0410259P.
XX
PR 12-SEP-2002; 2002US-0410260P.
XX

```
PA (INCY-) INCYTE CORP.
XX Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
PI Harthshorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirtson ES;
PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
PI Patury S, Shi X, Suarez CJ;
XX WPI; 2004-329368/30.
DR N-PSDB; ACN41271.
XX
XX New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT in diagnosing a condition, disease or disorder associated with human
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT in gene mapping.
XX
XX Claim 27; Page; 190pp; English.
XX
XX The invention relates to novel diagnostic and therapeutic polynucleotides
CC selected from one of the 2722 sequences defined in the specification. A
CC polynucleotide of the invention may have a use in gene therapy. The human
CC polynucleotide and therapeutic polynucleotides (dithp) or polypeptides may be
CC used to diagnose a particular condition, disease or disorder associated
CC with human molecules, e.g. cell proliferative disorders,
CC autoimmune/inflammatory disorder, developmental disorder, endocrine
CC disorder, neurological disorders, gastrointestinal disorders, or
CC infections caused by virus, bacteria, fungi or parasite. The dithp
CC molecules may also be used in genetic mapping, in identifying individuals
CC from minute biological samples, in detecting single nucleotide
CC polymorphisms, as molecular weight markers, and for somatic or germline
CC gene therapy. The present sequence represents a dithp protein of the
CC invention. Note: The sequence data for this patent is not represented in
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX
XX Sequence 254 AA;
XX
XX Query Match 24.6%; Score 376.5; DB 8; Length 254;
XX Best Local Similarity 30.2%; Pred. No. 1.7e-30;
XX Matches 83; Conservative 48; Mismatches 85; Indels 59; Gaps 4;
XX
Qy 24 GVRLLPPFCRGMEBIANFPVRPSDVWIVTPKSGTSLLOEVVYLVVSQADPDEIGLNMID 83
Db 17 GVPLIKYFAEALGPLQSFQARPDLLINTYPKSGTTWVSQILDIMYQGDLEKCNRAPIY 76
Qy 84 EQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPDLVVSYY 143
Db 77 VRVPFLE-----VVYVARNPKDVAVSYY 99
Qy 144 QFHRSLRTMSYRGTFQBFRCFRFMDKLGYSWFEHVQEFWEHRMDSNVLFPLKYEDMHRDL 203
Db 100 HFHRMEKAHPPEPGTWDSEFLKFMAGEYSYGSWYQHVQEWELSRTHPVLVLYFYEDMKENP 159
Qy 204 VTWVEQLARFLGVSCDKAQLAEALTEHC-----HQLVDQCCNAEALP-V 245
Db 160 KREIQKILEFVGRSLPEETMBDFMQHTSFKEMKKNPMTNITYTVQELMDHDSIS----PFM 215
Qy 246 GRGRVGLWKDIFTVSMNEKFLVTKQKMGKCDLTF 280
Db 216 RKGMAQDWKTTFTVAQNERFDADYAEKMGAGSLSF 250
XX
RESULT 91
ID ABM82617
XX ABM82617 standard; protein; 254 AA.
AC ABM82617;
XX
XX 18-NOV-2004 (first entry)
XX
```

```
DE (INCY-) INCYTE CORP.
XX Human diagnostic and therapeutic pprotein SEQ ID NO:2866.
XX gene therapy; human diagnostic and therapeutic polynucleotide; dithp.
XX Homo sapiens.
XX WO2004023973-A2.
XX 25-MAR-2004.
XX
XX 12-SEP-2003; 2003WO-US028227.
XX
XX 12-SEP-2002; 2002US-0410259P.
XX 12-SEP-2002; 2002US-0410260P.
XX (INCY-) INCYTE CORP.
XX
XX Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
PI Harthshorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirtson ES;
PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
PI Patury S, Shi X, Suarez CJ;
XX WPI; 2004-329368/30.
DR N-PSDB; ACN41269.
XX
XX New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT in diagnosing a condition, disease or disorder associated with human
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT in gene mapping.
XX
XX Claim 27; Page; 190pp; English.
XX
XX The invention relates to novel diagnostic and therapeutic polynucleotides
CC selected from one of the 2722 sequences defined in the specification. A
CC polynucleotide of the invention may have a use in gene therapy. The human
CC polynucleotide and therapeutic polynucleotides (dithp) or polypeptides may be
CC used to diagnose a particular condition, disease or disorder associated
CC with human molecules, e.g. cell proliferative disorders,
CC autoimmune/inflammatory disorder, developmental disorder, endocrine
CC disorder, neurological disorders, gastrointestinal disorders, or
CC infections caused by virus, bacteria, fungi or parasite. The dithp
CC molecules may also be used in genetic mapping, in identifying individuals
CC from minute biological samples, in detecting single nucleotide
CC polymorphisms, as molecular weight markers, and for somatic or germline
CC gene therapy. The present sequence represents a dithp protein of the
CC invention. Note: The sequence data for this patent is not represented in
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX
XX Sequence 254 AA;
XX
XX Query Match 24.6%; Score 376.5; DB 8; Length 254;
XX Best Local Similarity 30.2%; Pred. No. 1.7e-30;
XX Matches 83; Conservative 48; Mismatches 85; Indels 59; Gaps 4;
XX
Qy 24 GVRLLPPFCRGMEBIANFPVRPSDVWIVTPKSGTSLLOEVVYLVVSQADPDEIGLNMID 83
Db 17 GVPLIKYFAEALGPLQSFQARPDLLINTYPKSGTTWVSQILDIMYQGDLEKCNRAPIY 76
Qy 84 EQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPDLVVSYY 143
Db 77 VRVPFLE-----VVYVARNPKDVAVSYY 99
Qy 144 QFHRSLRTMSYRGTFQBFRCFRFMDKLGYSWFEHVQEFWEHRMDSNVLFPLKYEDMHRDL 203
Db 100 HFHRMEKAHPPEPGTWDSEFLKFMAGEYSYGSWYQHVQEWELSRTHPVLVLYFYEDMKENP 159
Qy 204 VTWVEQLARFLGVSCDKAQLAEALTEHC-----HQLVDQCCNAEALP-V 245
Db 100 HFHRMEKAHPPEPGTWDSEFLKFMAGEYSYGSWYQHVQEWELSRTHPVLVLYFYEDMKENP 159
Qy 204 VTWVEQLARFLGVSCDKAQLAEALTEHC-----HQLVDQCCNAEALP-V 245
XX
```



```
Db      160 KREIQKILEFVGRSLPBEETMDFMVQHTSFKEMKKNPMTNTYTTVPQELMDHSIS-----PFM 215
Qy      246 GRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db      216 RKGWAGDKTFTTVAQNERFDADYAEKMGAGCSLSF 250

RESULT 92
ABM82618
ID      ABM82618 standard; protein; 254 AA.
AC      ABM82618;
XX
DT      18-NOV-2004 (first entry)
XX
DE      Human diagnostic and therapeutic pprotein SEQ ID NO:2867.
XX
KW      gene therapy; human diagnostic and therapeutic polynucleotide; dithp.
XX
OS      Homo sapiens.
XX
FN      WO2004023973-A2.
XX
PD      25-MAR-2004.
XX
PF      12-SEP-2003; 2003WO-US028227.
XX
PR      12-SEP-2002; 2002US-0410259P.
PR      12-SEP-2002; 2002US-0410260P.
XX
PA      (INCY-) INCYTE CORP.
XX
PI      Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
PI      Harthorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
PI      Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;
PI      Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
PI      Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
PI      Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirtson ES;
PI      Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
PI      Patury S, Shi X, Suarez CJ;
XX
WPI; 2004-329368/30.
DR      N-PSDB; ACN41270.
XX
XX      New diagnostic and therapeutic polynucleotides and polypeptides, useful
XX      in diagnosing a condition, disease or disorder associated with human
XX      molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
XX      in gene mapping.
XX
PS      Claim 27; Page; 190pp; English.
XX
XX      The invention relates to novel diagnostic and therapeutic polynucleotides
XX      selected from one of the 2722 sequences defined in the specification. A
XX      polynucleotide of the invention may have a use in gene therapy. The human
XX      diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
XX      used to diagnose a particular condition, disease or disorder associated
XX      with human molecules, e.g. cell proliferative disorders,
XX      autoimmune/inflammatory disorder, developmental disorder, endocrine
XX      disorder, neurological disorders, gastrointestinal disorders, or
XX      infections caused by virus, bacteria, fungi or parasite. The dithp
XX      molecules may also be used in genetic mapping, in identifying individuals
XX      from minute biological samples, in detecting single nucleotide
XX      polymorphisms, as molecular weight markers, and for somatic or germline
XX      gene therapy. The present sequence represents a dithp protein of the
XX      invention. Note: The sequence data for this patent is not represented in
XX      the printed specification, but was obtained in electronic format directly
XX      from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX
SQ      Sequence 254 AA;
Query Match      24.6%; Score 376.5; DB 8; Length 254;
Best Local Similarity 30.2%; Pred. No. 1.7e-30;
Matches 83; Conservative 48; Mismatches 85; Indels 59; Gaps 4;
```

```
Qy      24 GVLPPFCRCGMEIEIANFVRPSDVMIIVTYPKSTSLQEVVYLVSGADPDEIGLMNID 83
Db      17 GVPLIKYFAEALGPLQSFQARPDLLINTYKSGTTWSQLDMYOGGLEKCNRAPIY 76
Qy      84 EQLPVLEYPQGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPDLVVSYY 143
Db      77 VRVPFLE-----VVVARNPKDVAVSYY 99
Qy      144 QFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHFVQEFWEHHRMDSNVFLFKYEDMRDL 203
Db      100 HFHRMEKAHPPEPGTWDSEFLKFMAGEVSYGSWYQHVQEWMLSRTHFVLYLFYEDMKENP 159
Qy      204 VTWVEQLARFLGVSCDKAQLEALTEHC-----HQLVDQCCNAEALP-V 245
Db      160 KREIQKILEFVGRSLPBEETMDFMVQHTSFKEMKKNPMTNTYTTVPQELMDHSIS-----PFM 215
Qy      246 GRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db      216 RKGWAGDKTFTTVAQNERFDADYAEKMGAGCSLSF 250

RESULT 93
ABM82620
ID      ABM82620 standard; protein; 254 AA.
XX
AC      ABM82620;
XX
DT      18-NOV-2004 (first entry)
XX
DE      Human diagnostic and therapeutic pprotein SEQ ID NO:2869.
XX
KW      gene therapy; human diagnostic and therapeutic polynucleotide; dithp.
XX
OS      Homo sapiens.
XX
FN      WO2004023973-A2.
XX
PD      25-MAR-2004.
XX
PF      12-SEP-2003; 2003WO-US028227.
XX
PR      12-SEP-2002; 2002US-0410259P.
PR      12-SEP-2002; 2002US-0410260P.
XX
PA      (INCY-) INCYTE CORP.
XX
PI      Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
PI      Harthorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
PI      Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;
PI      Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
PI      Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
PI      Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirtson ES;
PI      Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
PI      Patury S, Shi X, Suarez CJ;
XX
WPI; 2004-329368/30.
DR      N-PSDB; ACN41272.
XX
XX      New diagnostic and therapeutic polynucleotides and polypeptides, useful
XX      in diagnosing a condition, disease or disorder associated with human
XX      molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
XX      in gene mapping.
XX
PS      Claim 27; Page; 190pp; English.
XX
XX      The invention relates to novel diagnostic and therapeutic polynucleotides
XX      selected from one of the 2722 sequences defined in the specification. A
XX      polynucleotide of the invention may have a use in gene therapy. The human
XX      diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
XX      used to diagnose a particular condition, disease or disorder associated
XX      with human molecules, e.g. cell proliferative disorders,
XX      autoimmune/inflammatory disorder, developmental disorder, endocrine
XX      disorder, neurological disorders, gastrointestinal disorders, or
XX      infections caused by virus, bacteria, fungi or parasite. The dithp
XX      molecules may also be used in genetic mapping, in identifying individuals
XX      from minute biological samples, in detecting single nucleotide
XX      polymorphisms, as molecular weight markers, and for somatic or germline
XX      gene therapy. The present sequence represents a dithp protein of the
XX      invention. Note: The sequence data for this patent is not represented in
XX      the printed specification, but was obtained in electronic format directly
XX      from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX
SQ      Sequence 254 AA;
```

CC disorder, neurological disorders, gastrointestinal disorders, or
CC infections caused by virus, bacteria, fungi or parasite. The dithp
CC molecules may also be used in genetic mapping, in identifying individuals
CC from minute biological samples, in detecting single nucleotide
CC polymorphisms, as molecular weight markers, and for somatic or germline
CC gene therapy. The present sequence represents a dithp protein of the
CC invention. Note: The sequence data for this patent is not represented in
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX
SQ Sequence 254 AA;

Query Match 24.6%; Score 376.5; DB 8; Length 254;
Best Local Similarity 30.2%; Pred. No. 1.7e-30;
Matches 83; Conservative 48; Mismatches 85; Indels 59; Gaps 4;
QY 24 GYRLPPFCRGKMEETANPPRPSDVWIVTPYKSGTSLLOEVVLYVLSQGDADPEIGLMNID 83
DB 17 GVPLIKYFAEALGPQSFQARPDLLINTYPSGTTWVSQILDMYQGGDLKCNRAPIY 76
QY 84 EQLPVLEYPPQGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPDLVWSY 143
DB 77 VRVPFLE-----VYTVARNPKDAVSY 99
QY 144 QHRSRLRTWSYRGCTQFPCRRFMNDKLGYSWFHEVQEFWEHRMDSNVLFKYEDMHRDL 203
DB 100 HFHRMEKAHPERGTWDSFLEKFMAGEVSYGSYQHVQEWELSRTHPLVLYLFYEDMKENP 159
QY 204 VTMVQLARFLGVSCDKAQLAETHC-----HQLVDCCNAEALP-V 245
DB 160 KREIQKLEFVGRSLPEETMDFWQHTSPFKEMKNPMTNTYTPQELMDHSIS----PPM 215
QY 246 GRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
DB 216 RKGMDGDKNTTFTVAQNERFDADYAEKMGACSLSF 250

RESULT 94
ADG76311
ID ADG76311 standard; protein; 259 AA.
XX
AC ADG76311;
XX
DT 11-MAR-2004 (first entry)
XX
DE Human Incyte enzymatic protein (SeqID 3).
XX
KW enzyme; human; biogenesis; biodegradation; autoimmune; inflammatory;
KW infectious; neurological; cardiovascular; metabolic; cell proliferative;
KW cancer; cytostatic; immunomodulator.
XX
OS Homo sapiens.
XX
PN WO2003083082-A2.
XX
PD 09-OCT-2003.
XX
PF 26-MAR-2003; 2003WO-US009772.
XX
PR 29-MAR-2002; 2002US-0368721P.
PR 29-MAR-2002; 2002US-0368799P.
XX
PA (INCY-) INCYTE CORP.
XX
PI Tran UK, Kable AF, Ramkumar J, Ison CH, Richardson TW, Lee SY;
PI Khare R, Marquis JP, Swarnakar A, Chawla NK, Elliott VS;
PI Emerling BM, Becha SD, Hafalia AJA, Li JX, Griffin JA, Hawkins PR;
PI Jin P, Chien D, Jiang X, Jackson AA, Mason PM, Bhatia UG;
PI Burrill JD, Lee S, Blake JJ, Ho A, Zheng W, Gao J;
XX
DR WPI; 2003-804054/75.
DR N-PSDB; ADG76366.
XX

PT New human enzymes (ENZM) polypeptide, useful for preparing a composition
PT for treating a disease associated with decreased expression or
PT overexpression of ENZM e.g. cancer.
XX
PS Claim 58; SEQ ID NO 3; 385pp; English.
XX
CC This invention relates to novel isolated polynucleotides, and the enzymes
CC encoded thereof, as well as any biologically active or immunogenic
CC fragments. Specifically, it refers to human enzymes involved in the
CC cellular processes of biogenesis and biodegradation, for example
CC oxidoreductases, hydrolases and isomerases. The present invention
CC describes the use of these nucleic acids and proteins (including single
CC nucleotide polymorphisms) in the diagnosis, treatment and prevention of
CC various disorders such as autoimmune/ inflammatory, infectious,
CC neurological, cardiovascular, metabolic and cell proliferative such as
CC cancer. Furthermore, the polypeptide is useful for preparing a
CC composition for treating a disease associated with decreased expression
CC or overexpression of the functional enzyme (ENZM) e.g. immune disorders.
CC Accordingly, the compositions exhibit cytostatic and immunomodulator
CC activities. This polypeptide sequence is a human incyte protein, an
CC enzyme of the invention.
XX
SQ Sequence 259 AA;

Query Match 23.8%; Score 363.5; DB 7; Length 259;
Best Local Similarity 29.6%; Pred. No. 4e-29;
Matches 83; Conservative 53; Mismatches 91; Indels 53; Gaps 5;
QY 18 KYVEFHGVLRLPPFCRGKMEETANPPRPSDVWIVTPYKSGTSLLOEVVLYVLSQGDADPEI 77
DB 12 KXVEGELLQAPVVDNWSQIQSFQAKPDLLICTYPRAGTTWIOEIVDMIEQNGDVEKC 71
QY 78 GLMNIDEQLPVLEYPPQGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPDK 137
DB 72 QRATIQHRHPFIEWARP-----PQ-----PS-----ARNAKD 98
QY 138 LVVSYQFHSRLRTWSYRGCTQFPCRRFMNDKLGYSWFHEVQEFWEHRMDSNVLFKYE 197
DB 99 CMVSYHFFQRMNHLMPDEGTWEEYFETFINGKVVWGSWFDHVKGWEMKDRHOILFLFY 158
QY 198 DMHRDLVTMVQLARFLGVSCDKAQLAETHC-----CHQLVDCCNA 240
DB 159 DIKDPKHEIRKVMQFMGKKVDETVLDKIVQETSEKMKENPMTNRTSVTSKILDSISS 218
QY 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
DB 219 F---MRKCTVGDWKNHFTVAQNERFDEIYRRKMGCTSNF 255

RESULT 95
ABB61517
ID ABB61517 standard; protein; 338 AA.
XX
AC ABB61517;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 11343.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX

PA (PEKE) PE CORP NY.
 XX Venter JC, Adams M, Li PWD, Myers EW;
 FI WPI; 2001-656860/75.
 XX N-PSDB; ABL05620.
 DR
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions.
 XX
 PS Disclosure; SEQ ID NO 11343; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 338 AA;
 Query Match 23.6%; Score 360.5; DB 4; Length 338;
 Best Local Similarity 31.0%; Pred. No. 1.2e-28;
 Matches 89; Conservative 51; Mismatches 112; Indels 35; Gaps 7;
 QY 9 PSTP--GFESKYFFHGVRLPPFCRGKMEETANFPVPSDVWIVTPKSGTSLLOEVVY 66
 DB 34 PSVPVVGWNEQRF-----CRLADTQPLDRVDFEVRDDWIVTLPKCGTTWQELAW 88
 QY 67 LVSGQADPDEIGLNMIDQLPVLEY-----PQGLDIIKELTSPRIKSHLPYRFLPS 119
 DB 89 LVINECDPETAASVDLTHRSPLFEGVVPVPHDTIAAANALPSRLIKSHLPAMWLP 148
 QY 120 DLHNGDSKVIYMARPKDLVSYQYFHRSLRTMSYRGTFQFCRPFMDKLGYSWFH 179
 DB 149 QIWSKRPKIIVYRNPKDAISYFFHWRGM--VGQGTKSDPMHGFIDGYVNFPCWPHI 206
 QY 180 QBFWEHRMDSNVLFUKYEDMHRDLVTWVEQLARFLGVSCKAQLEALTEHC--HOLVQ- 236
 DB 207 LDFWQLRHEPNIFFTSYERMKQLGQVISEVAQFLERSVQEQMQMQORHLSFESMRNP 266
 QY 237 CCN-----AEALPVGRGRVGLWKDIFTVSMNEKFDL 267
 DB 267 ACNHVKFESMKAAAGREVEEPFRVRRGVVGSHKDELTAIIRFDL 313

RESULT 96
 ABB62612
 ID ABB62612 standard; protein; 346 AA.
 AC ABB62612;
 XX
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 14628.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US009231.
 XX
 PF 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.
 XX (PEKE) PE CORP NY.
 XX Venter JC, Adams M, Li PWD, Myers EW;
 FI WPI; 2001-656860/75.
 XX N-PSDB; ABL06715.
 DR
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions.
 XX
 PS Disclosure; SEQ ID NO 14628; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 346 AA;
 Query Match 23.5%; Score 359.5; DB 4; Length 346;
 Best Local Similarity 32.3%; Pred. No. 1.6e-28;
 Matches 82; Conservative 35; Mismatches 88; Indels 49; Gaps 5;
 QY 22 FHGVR-----LPPFCRGKMEETANFPVPSDVWIVTPKSGTSLLOEVVY 69
 DB 27 FHGERTQVQVSGYGFPHKYKDEAERYNFARPDWVATVPRSGTTWTQELIWLVA 86
 QY 70 QGADPDEIGLNMIDQLPVLEY-----PQGLDIIK 100
 DB 87 NGLDFEHAQERPLTERPFFPFFLVHPKIKEELQENRDSAEALFEKTIARFGYEALS 146
 QY 101 EL--TSPRLIKSHLPYRFLPSDLHNGDSKVIYMARPNKDLVSYQYFHRSLRTMSYRGTF 158
 DB 147 EIPRSQRRFIKTHFPFSLMPPSVLEKKCKVIVVVRDPKDVAVSYHNLRLFRFTQGVGDF 206
 QY 159 QFCRCRRNDKLGYSW---FEHVQEFWEHRMDSNVLFUKYEDMHRDLVTWVEQLARFLG 215
 DB 207 ERYWHYFQN---GLNPLPYSHVKEAREHAHLSNVLFRLRYEDMLADLPGAINSTASPLE 263
 QY 216 VSCDKAOLEALTEH 229
 DB 264 CFPKPEDMDRLLDH 277
 RESULT 97
 ABB62695
 ID ABB62695 standard; protein; 254 AA.
 XX
 XX
 AC ABB62695;
 XX
 DT 18-NOV-2004 (first entry)
 XX
 DE Human diagnostic and therapeutic pprotein SEQ ID NO:2944.
 XX
 KW gene therapy; human diagnostic and therapeutic polynucleotide; dithp.
 XX
 OS Homo sapiens.
 XX
 PN WO2004023973-A2.
 XX
 PD 25-MAR-2004.
 XX
 PF 12-SEP-2003; 2003WO-US028227.
 XX

```
PR 12-SEP-2002; 2002US-0410259P.
XX 12-SEP-2002; 2002US-0410260P.
XX (INCY-) INCYTE CORP.
XX
XX Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
PI Harthshorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;
PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
PI Patury S, Shi X, Suarez CJ;
XX
XX WPI; 2004-329368/30.
DR N-PSDB; ACN41347.
XX
XX New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT in diagnosing a condition, disease or disorder associated with human
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT in gene mapping.
XX
XX Claim 27; Page; 190pp; English.
XX
XX The invention relates to novel diagnostic and therapeutic polynucleotides
CC selected from one of the 2722 sequences defined in the specification. A
CC polynucleotide of the invention may have a use in gene therapy. The human
CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
CC used to diagnose a particular condition, disease or disorder associated
CC with human molecules, e.g. cell proliferative disorders,
CC autoimmune/inflammatory disorder, developmental disorders, endocrine
CC disorder, neurological disorders, gastrointestinal disorders, or
CC infections caused by virus, bacteria, fungi or parasite. The dithp
CC molecules may also be used in genetic mapping, in identifying individuals
CC from minute biological samples, in detecting single nucleotide
CC polymorphisms, as molecular weight markers, and for somatic or germline
CC gene therapy. The present sequence represents a dithp protein of the
CC invention. Note: The sequence data for this patent is not represented in
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX
XX Sequence 254 AA;
SQ
Query Match 23.4%; Score 357.5; DB 8; Length 254;
Best Local Similarity 29.9%; Pred. No. 1.7e-28;
Matches 81; Conservative 46; Mismatches 93; Indels 51; Gaps 4;
QY 24 GVRLPFFCRGKMEETANFVRPSDVWIVTYPKSGTSLLOEVVLYVSQGADPDEIGLMNID 83
DB 17 GVPLIKYFAEALGPLQSFQARPDDLITSTYPKSGTTWVSQILDMTYQGGD----- 66
QY 84 EQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVYIMARNPKDLVWSY 143
DB 67 -----LEKCHR-----VPQTLDDQKVKVYVARNAKDVAVSYY 99
QY 144 QFHRSLRTMSYRGTFQFCRCRPMNDKLGYSWFEHVOEFWEHRMDSNVFLFKYEDMHRDL 203
DB 100 HFYHMAKVHPPEFGTWDSFLEKFMVGEVSGYSWYQHVQEWELSRTHPVLVLYFYEDMKENP 159
QY 204 VTMVQLARFLVSCDKAQLAETHC---HQLVQCCNAEALP-----VGRGR 249
DB 160 KREIQKILEFVGRSLPEETVDFVVOHTSPFKEMKKNPMTNTYTVQFEMDHSISPFMRKGM 219
QY 250 VGLWKDIFTVMNKEKFDLYVKMGKCDLTF 280
DB 220 AGDWKTFTTVAQNERFDADYAEKMACCSLSF 250
RESULT 98
ABM82697
ID ABM82697 standard; protein; 254 AA.
XX
AC ABM82697;
```

```
XX 18-NOV-2004 (first entry)
XX Human diagnostic and therapeutic pprotein SEQ ID NO:2946.
XX
XX gene therapy; human diagnostic and therapeutic polynucleotide; dithp.
XX
XX Homo sapiens.
XX
XX WO2004023973-A2.
XX
XX 25-MAR-2004.
XX
XX 12-SEP-2003; 2003WO-US028227.
XX
XX 12-SEP-2002; 2002US-0410259P.
XX 12-SEP-2002; 2002US-0410260P.
XX (INCY-) INCYTE CORP.
XX
XX Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
PI Harthshorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;
PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
PI Patury S, Shi X, Suarez CJ;
XX
XX WPI; 2004-329368/30.
DR N-PSDB; ACN41349.
XX
XX New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT in diagnosing a condition, disease or disorder associated with human
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT in gene mapping.
XX
XX Claim 27; Page; 190pp; English.
XX
XX The invention relates to novel diagnostic and therapeutic polynucleotides
CC selected from one of the 2722 sequences defined in the specification. A
CC polynucleotide of the invention may have a use in gene therapy. The human
CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
CC used to diagnose a particular condition, disease or disorder associated
CC with human molecules, e.g. cell proliferative disorders,
CC autoimmune/inflammatory disorder, developmental disorders, endocrine
CC disorder, neurological disorders, gastrointestinal disorders, or
CC infections caused by virus, bacteria, fungi or parasite. The dithp
CC molecules may also be used in genetic mapping, in identifying individuals
CC from minute biological samples, in detecting single nucleotide
CC polymorphisms, as molecular weight markers, and for somatic or germline
CC gene therapy. The present sequence represents a dithp protein of the
CC invention. Note: The sequence data for this patent is not represented in
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX
XX Sequence 254 AA;
SQ
Query Match 23.4%; Score 357.5; DB 8; Length 254;
Best Local Similarity 29.9%; Pred. No. 1.7e-28;
Matches 81; Conservative 46; Mismatches 93; Indels 51; Gaps 4;
QY 24 GVRLPFFCRGKMEETANFVRPSDVWIVTYPKSGTSLLOEVVLYVSQGADPDEIGLMNID 83
DB 17 GVPLIKYFAEALGPLQSFQARPDDLITSTYPKSGTTWVSQILDMTYQGGD----- 66
QY 84 EQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVYIMARNPKDLVWSY 143
DB 67 -----LEKCHR-----VPQTLDDQKVKVYVARNAKDVAVSYY 99
QY 144 QFHRSLRTMSYRGTFQFCRCRPMNDKLGYSWFEHVOEFWEHRMDSNVFLFKYEDMHRDL 203
DB 100 HFYHMAKVHPPEFGTWDSFLEKFMVGEVSGYSWYQHVQEWELSRTHPVLVLYFYEDMKENP 159
```

QY 204 VTMVEQLARFLGVSCDKAQLALETHC---HQLVDQCCNAELP-----VGRGR 249
Db 160 KREIQKILEFVGRSLFEETVDFVVOHTSFKEKMKPNMTNTTVPQEFMDHSISPPMRKGM 219
QY 250 VGLWKDIFTVSNMNEKFDLVYKQMKGKCDLTF 280
Db 220 AGDWKTTTVAQNERFDADYAEKMGCSLSF 250
RESULT 99
ID ADD48500 standard; protein; 283 AA.
XX
AC ADD48500;
XX
XX 02-DEC-2004 (revised)
DT 29-JAN-2004 (first entry)
XX
XX Rat Protein CAA45007, SEQ ID NO 14201.
XX
XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
XX Rattus norvegicus.
OS Unidentified.
XX
XX WO2003016475-A2.
PN
XX
XX 27-FEB-2003.
PD
XX
XX 14-AUG-2002; 2002WO-US025765.
PF
XX
XX 14-AUG-2001; 2001US-0312147P.
PR
XX 01-NOV-2001; 2001US-0346382P.
PR
XX 26-NOV-2001; 2001US-0333347P.
XX
XX (GCHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX
XX WPI; 2003-268312/26.
DR
XX GENBANK; CAA45007.
XX
PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
XX Example 1; Page; 1017pp; English.
PS
XX
CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (described in Table 3

CC of the specification) which is differentially expressed during pain.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 283 AA;

Query Match 23.4%; Score 357.5; DB 7; Length 283;
Best Local Similarity 33.5%; Pred. No. 1.9e-28;
Matches 92; Conservative 53; Mismatches 105; Indels 25; Gaps 8;

QY 19 YFEFHGVRLLPPFCRGK--MEBIAN-FPVRPDSVMIIVTPKSGTSLLOQVWVYLVSGADPD 75
Db 3 YTFEGIPFPFAFGIPKETLQNVCKNFVVKBEDLLLTYPKSGTNWLIIEIVCLIQTKGDPK 62
QY 76 EIGLWNIDEQLPVLEYPOPGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNP 135
Db 63 WIQSVTIWDRSPWIE-TDLGYDMLIKKKGPRLITSHLPMLHLSKLSFKAKVIYLIRNP 121
QY 136 KDLVVYVYQFHRSLRTMSYR---GTFQEFCCRFRMNDKLGYGSMFHEHVOEFWEHRMDSNV 191
Db 122 RDVLVSGYFWM-GKTTLAKPDSLGTVVEM---FLKGVYPYGSWFHEHRAWLSMRELDNF 177
QY 192 LFLKYEDMHRDLVTMVEQLARFLGVSCDKAQLALETHCHQLVDQCCNAE-----A 242
Db 178 LLLYEDMKDVTMTGTTIKKICDFLCKLEPDELVLKYSSFQVMKENNMYNLMWEKELI 237
QY 243 LP---VGRGRVGLWKDIFTVSNMNEKFDLVYKQMK 273
Db 238 LPGAFTFRNGTTGDWKNHFTVAQAEAFDKVFOEKM 272

RESULT 100

AEF05774
ID AEF05774 standard; protein; 285 AA.

XX
AC AEF05774;

XX
DT 23-FEB-2006 (first entry)

XX
DE Human hydroxysteroid sulfotransferase.

XX
KW drug screening; metabolism; DNA typing;

XX
KW hydroxysteroid sulfotransferase 2A1; SULT2A1; enzyme.

XX
OS Homo sapiens.

XX
PN WO2005123922-A1.

XX
PD 29-DEC-2005.

XX
PF 17-JUN-2005; 2005WO-CA000941.

XX
PR 17-JUN-2004; 2004US-0580540P.

XX
PA (UYGU-) UNIV GUELPH.

XX
PI Squires J, Sinclair P;

XX
XX WPI; 2006-067510/07.

XX
PT New polynucleotides, useful in screening for a substance that enhances
PT metabolism of 16-androstene steroids or in reducing or preventing boar
PT taint in a pig.

XX
PS Disclosure; Fig 5; 93pp; English.

XX
CC The new invention relates to a polynucleotide encoding a
CC sulfotransferase. The sequence comprises a fully defined 1036-bp sequence
CC (SEQ ID NO: 1); sequence encoding a protein having a fully defined 286-
CC amino acid sequence (SEQ ID NO: 2); a sequence encoding a
CC sulfotransferase that hybridizes to the sequence in (A); or a sequence
CC having at least about 90% identity with the sequences in (A). Also

CC described are a vector comprising the polynucleotide operatively linked
CC to a regulatory sequence; a recombinant host cell comprising the vector;
CC a polypeptide; an antibody that specifically binds to the polypeptide; a
CC method of determining the genetic capacity of a male pig to develop boar
CC taint; a method for reducing or preventing boar taint in a pig; a method
CC for screening for a substance that enhances metabolism of 16-androstene
CC steroids in a pig; a method for screening for a substance that enhance
CC metabolism of 16-androstene steroids by enhancing transcription and/or
CC translation of the nucleotide sequence encoding sulfotransferase 2A1
CC (SULT2A1); and a method of genetically typing pigs to determine those
CC with desired boar taint characteristics. The isolated polynucleotide
CC comprises a nucleotide change of a cytosine (C) to a thymine (T) at
CC nucleotide position 219. The polypeptide is a SULT2A1 hydroxysteroid
CC sulfotransferase. The polynucleotide is useful in screening for a
CC substance that enhances metabolism of 16-androstene steroids or in
CC reducing or preventing boar taint in a pig. The present sequence is a
CC human hydroxysteroid sulfotransferase variant.

XX SQ Sequence 285 AA;

Query Match	23.3%;	Score 356.5;	DB 10;	Length 285;
Best Local Similarity	31.0%;	Pred. No. 2.5e-28;		
Matches	83;	Conservative 59;	Mismatches 109;	Indels 17; Gaps 5;
QY	22	FHGVRLLP--FCRGKMEETAN--PPVRPSDVWIVTPKSGTSLQVWVYLVSQGADPDEIG	78	
Db	8	FGIAFPTMGFRSETLRKVRDEFVIRDEVDVILITPKSGTNWLABILCLMHSGDAKWITQ	67	
QY	79	LMNIDEQLPVLEYPQGLDIIKELTSPRLIKSHLPYRFLPDLHNGDSKVIYMARNP KD L	138	
Db	68	SVPIWERSFWVE-SGIGYTALSETSPRLFSSHLPIQLFPKSFPSKAKVIYLMENPRDV	126	
QY	139	VVSYQFHRSKLTMSYRGTFQBFCCRFMNDKLGYSWFHVOEFWEHRMDSNVLFKYED	198	
Db	127	LVSGYFFWKNMKFIKKPKSWEEYFEWFCQGTVVYGSWFDHIGWMPMREKNFLLSYEE	186	
QY	199	MHRDLVTWVEQLARFLGVSCDKAQLALETHC---HQLVDQCCNAEALPVG-----	246	
Db	187	LKQDTGRTIEKICQFLGKLTLEPEELNLILKNSSFQSMKENKMSNYSLLSVDYVVDKAQLL	246	
QY	247	-RGRVGLWKDIFTVSMNEKFDLVYKOKM	273	
Db	247	RKGVSGDNKNHFTVAQAEFDKLFQEKM	274	

Search completed: May 17, 2006, 10:59:05
Job time : 211 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 17, 2006, 11:05:51 ; Search time 601 Seconds
(without alignments)
721.557 Million cell updates/sec

Title: US-10-768-158-2

Perfect score: 1530

Sequence: 1 MAESAEPTSPGPEFSKYF.....FDLVYKQMGKCDLTFDFYL 284

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 8366291 seqs, 1526956180 residues

Total number of hits satisfying chosen parameters: 8366291

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Pending Patents_AA_Main:*

1: /EMC_Celerra_SIDS3/ptodata/2/paa/PCTUS_COMB.pep.*
2: /EMC_Celerra_SIDS3/ptodata/2/paa/US066_COMB.pep.*
3: /EMC_Celerra_SIDS3/ptodata/2/paa/US073_COMB.pep.*
4: /EMC_Celerra_SIDS3/ptodata/2/paa/US074_COMB.pep.*
5: /EMC_Celerra_SIDS3/ptodata/2/paa/US075_COMB.pep.*
6: /EMC_Celerra_SIDS3/ptodata/2/paa/US076_COMB.pep.*
7: /EMC_Celerra_SIDS3/ptodata/2/paa/US077_COMB.pep.*
8: /EMC_Celerra_SIDS3/ptodata/2/paa/US078_COMB.pep.*
9: /EMC_Celerra_SIDS3/ptodata/2/paa/US079_COMB.pep.*
10: /EMC_Celerra_SIDS3/ptodata/2/paa/US080_COMB.pep.*
11: /EMC_Celerra_SIDS3/ptodata/2/paa/US081_COMB.pep.*
12: /EMC_Celerra_SIDS3/ptodata/2/paa/US082_COMB.pep.*
13: /EMC_Celerra_SIDS3/ptodata/2/paa/US083_COMB.pep.*
14: /EMC_Celerra_SIDS3/ptodata/2/paa/US084_COMB.pep.*
15: /EMC_Celerra_SIDS3/ptodata/2/paa/US085_COMB.pep.*
16: /EMC_Celerra_SIDS3/ptodata/2/paa/US086_COMB.pep.*
17: /EMC_Celerra_SIDS3/ptodata/2/paa/US087_COMB.pep.*
18: /EMC_Celerra_SIDS3/ptodata/2/paa/US088_COMB.pep.*
19: /EMC_Celerra_SIDS3/ptodata/2/paa/US089_COMB.pep.*
20: /EMC_Celerra_SIDS3/ptodata/2/paa/US090_COMB.pep.*
21: /EMC_Celerra_SIDS3/ptodata/2/paa/US091_COMB.pep.*
22: /EMC_Celerra_SIDS3/ptodata/2/paa/US092_COMB.pep.*
23: /EMC_Celerra_SIDS3/ptodata/2/paa/US093_COMB.pep.*
24: /EMC_Celerra_SIDS3/ptodata/2/paa/US094_COMB.pep.*
25: /EMC_Celerra_SIDS3/ptodata/2/paa/US095_COMB.pep.*
26: /EMC_Celerra_SIDS3/ptodata/2/paa/US096_COMB.pep.*
27: /EMC_Celerra_SIDS3/ptodata/2/paa/US097_COMB.pep.*
28: /EMC_Celerra_SIDS3/ptodata/2/paa/US098_COMB.pep.*
29: /EMC_Celerra_SIDS3/ptodata/2/paa/US099_COMB.pep.*
30: /EMC_Celerra_SIDS3/ptodata/2/paa/US100_COMB.pep.*
31: /EMC_Celerra_SIDS3/ptodata/2/paa/US101_COMB.pep.*
32: /EMC_Celerra_SIDS3/ptodata/2/paa/US102_COMB.pep.*
33: /EMC_Celerra_SIDS3/ptodata/2/paa/US103_COMB.pep.*
34: /EMC_Celerra_SIDS3/ptodata/2/paa/US104_COMB.pep.*
35: /EMC_Celerra_SIDS3/ptodata/2/paa/US105_COMB.pep.*
36: /EMC_Celerra_SIDS3/ptodata/2/paa/US106_COMB.pep.*
37: /EMC_Celerra_SIDS3/ptodata/2/paa/US107_COMB.pep.*
38: /EMC_Celerra_SIDS3/ptodata/2/paa/US108_COMB.pep.*
39: /EMC_Celerra_SIDS3/ptodata/2/paa/US109_COMB.pep.*
40: /EMC_Celerra_SIDS3/ptodata/2/paa/US110_COMB.pep.*
41: /EMC_Celerra_SIDS3/ptodata/2/paa/US111_COMB.pep.*
42: /EMC_Celerra_SIDS3/ptodata/2/paa/US112_COMB.pep.*
43: /EMC_Celerra_SIDS3/ptodata/2/paa/US113_COMB.pep.*

44: /EMC_Celerra_SIDS3/ptodata/2/paa/US114_COMB.pep.*
45: /EMC_Celerra_SIDS3/ptodata/2/paa/US600_COMB.pep.*
46: /EMC_Celerra_SIDS3/ptodata/2/paa/US601_COMB.pep.*
47: /EMC_Celerra_SIDS3/ptodata/2/paa/US602_COMB.pep.*
48: /EMC_Celerra_SIDS3/ptodata/2/paa/US603_COMB.pep.*
49: /EMC_Celerra_SIDS3/ptodata/2/paa/US604_COMB.pep.*
50: /EMC_Celerra_SIDS3/ptodata/2/paa/US605_COMB.pep.*
51: /EMC_Celerra_SIDS3/ptodata/2/paa/US606_COMB.pep.*
52: /EMC_Celerra_SIDS3/ptodata/2/paa/US607_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1530	100.0	284	1	PCT-US01-41986-4	Sequence 4, Appli
2	1530	100.0	284	1	PCT-US01-41986-6	Sequence 6, Appli
3	1530	100.0	284	1	PCT-US02-11654-60	Sequence 60, Appl
4	1530	100.0	284	1	PCT-US02-11654-81	Sequence 81, Appl
5	1530	100.0	284	1	PCT-US02-11654-103	Sequence 103, App
6	1530	100.0	284	1	PCT-US02-11654-104	Sequence 104, App
7	1530	100.0	284	1	PCT-US04-0750-110	Sequence 110, App
8	1530	100.0	284	1	PCT-US04-02851-2	Sequence 2, Appli
9	1530	100.0	284	27	US-09-791-537-17119	Sequence 17119, A
10	1530	100.0	284	31	US-10-121-024B-60	Sequence 60, Appl
11	1530	100.0	284	31	US-10-121-024B-81	Sequence 81, Appl
12	1530	100.0	284	31	US-10-121-024B-103	Sequence 103, App
13	1530	100.0	284	31	US-10-121-024B-104	Sequence 104, App
14	1530	100.0	284	31	US-10-121-024B-708	Sequence 708, App
15	1530	100.0	284	31	US-10-121-024C-60	Sequence 60, Appl
16	1530	100.0	284	31	US-10-121-024C-81	Sequence 81, Appl
17	1530	100.0	284	31	US-10-121-024C-103	Sequence 103, App
18	1530	100.0	284	31	US-10-121-024C-104	Sequence 104, App
19	1530	100.0	284	31	US-10-121-024C-708	Sequence 708, App
20	1530	100.0	284	34	US-10-427-631-7	Sequence 7, Appli
21	1530	100.0	284	35	US-10-573-989-17	Sequence 17, Appl
22	1530	100.0	284	37	US-10-757-262-110	Sequence 110, App
23	1530	100.0	284	37	US-10-768-158-2	Sequence 2, Appli
24	1530	100.0	284	39	US-10-932-348-408	Sequence 408, App
25	1530	100.0	284	39	US-10-990-328-7686	Sequence 7686, Ap
26	1530	100.0	284	39	US-10-990-328A-7686	Sequence 7686, Ap
27	1530	100.0	284	41	US-11-154-298-60	Sequence 60, Appl
28	1530	100.0	284	41	US-11-154-298-81	Sequence 81, Appl
29	1530	100.0	284	41	US-11-154-298-103	Sequence 103, App
30	1530	100.0	284	41	US-11-154-298-104	Sequence 104, App
31	1530	100.0	284	41	US-11-154-298-708	Sequence 708, App
32	1530	100.0	284	41	US-11-155-906-60	Sequence 60, Appl
33	1530	100.0	284	41	US-11-155-906-81	Sequence 81, Appl
34	1530	100.0	284	41	US-11-155-906-103	Sequence 103, App
35	1530	100.0	284	41	US-11-155-906-104	Sequence 104, App
36	1530	100.0	284	41	US-11-155-906-708	Sequence 708, App
37	1530	100.0	284	41	US-11-156-231-60	Sequence 60, Appl
38	1530	100.0	284	41	US-11-156-231-81	Sequence 81, Appl
39	1530	100.0	284	41	US-11-156-231-103	Sequence 103, App
40	1530	100.0	284	41	US-11-156-231-104	Sequence 104, App
41	1530	100.0	284	41	US-11-156-231-708	Sequence 708, App
42	1530	100.0	284	49	US-60-444-781-2	Sequence 2, Appli
43	1530	100.0	284	49	US-60-491-156-24	Sequence 24, Appl
44	1530	100.0	284	49	US-60-499-964-408	Sequence 408, App
45	1530	100.0	284	49	US-10-170-205E-16293X	Sequence 16293, A
46	1530	100.0	311	47	US-60-245-241-166	Sequence 166, App
47	1530	100.0	311	47	US-60-258-016-46	Sequence 46, Appl
48	1506	98.4	284	27	US-09-791-537-89260	Sequence 89260, A
49	1506	98.4	284	32	US-10-205-331-46	Sequence 46, Appl
50	1506	98.4	284	32	US-10-205-331-46	Sequence 225, App
51	1502.5	98.2	319	47	US-60-229-512-225	Sequence 27511, A
52	1332	87.1	260	31	US-10-170-205E-27511	Sequence 406, App
53	1332	87.1	260	39	US-10-932-348-406	Sequence 7684, App
54	1332	87.1	260	39	US-10-990-328-7684	

55	1332	87.1	260	39	US-10-990-328A-7684	Sequence 7684, Ap
56	1332	87.1	260	49	US-60-499-964-406	Sequence 406, App
57	1267	82.8	286	39	US-10-932-348-407	Sequence 407, App
58	1267	82.8	286	39	US-10-990-328-7685	Sequence 7685, Ap
59	1267	82.8	286	39	US-10-990-328A-7685	Sequence 7685, Ap
60	1267	82.8	286	49	US-60-499-964-407	Sequence 407, App
61	1002	65.5	218	1	PCT-US02-11654-172	Sequence 172, App
62	1002	65.5	218	31	US-10-121-024B-172	Sequence 172, App
63	1002	65.5	218	31	US-10-121-024C-172	Sequence 172, App
64	1002	65.5	218	41	US-11-154-298-172	Sequence 172, App
65	1002	65.5	218	41	US-11-155-906-172	Sequence 172, App
66	1002	65.5	218	41	US-11-156-231-172	Sequence 172, App
67	867.5	56.7	171	1	PCT-US02-11654-173	Sequence 173, App
68	867.5	56.7	171	31	US-10-121-024B-173	Sequence 173, App
69	867.5	56.7	171	31	US-10-121-024C-173	Sequence 173, App
70	867.5	56.7	171	35	US-10-573-989-18	Sequence 18, Appl
71	867.5	56.7	171	41	US-11-154-298-173	Sequence 173, App
72	867.5	56.7	171	41	US-11-155-906-173	Sequence 173, App
73	867.5	56.7	171	41	US-11-156-231-173	Sequence 173, App
74	860	56.2	160	1	PCT-US01-41986-2	Sequence 2, Appli
75	804	52.5	194	31	US-10-170-205E-27771	Sequence 27771, A
76	495	32.4	298	27	US-09-791-537-24613	Sequence 24613, A
77	495	32.4	298	27	US-09-791-537-26347	Sequence 26347, A
78	489	32.0	291	27	US-09-791-537-66629	Sequence 66629, A
79	489	31.4	296	30	US-10-072-012-667	Sequence 667, App
80	480	31.4	592	47	US-60-212-357-126	Sequence 126, App
81	480	31.4	592	47	US-60-229-512-201	Sequence 201, App
82	479	31.3	295	27	US-09-791-537-52308	Sequence 52308, A
83	474	31.0	1305	1	PCT-US01-08631-57913	Sequence 57913, A
84	474	31.0	1305	34	US-10-450-763-57913	Sequence 57913, A
85	472	30.8	295	1	PCT-US01-41986-17	Sequence 17, Appl
86	472	30.8	295	1	PCT-US05-01474-4	Sequence 4, Appli
87	472	30.8	295	14	US-08-484-878-7	Sequence 7, Appli
88	472	30.8	295	27	US-09-791-537-67010	Sequence 67010, A
89	472	30.8	295	31	US-10-170-205E-29903	Sequence 29903, A
90	472	30.8	295	31	US-10-170-205E-29904	Sequence 29904, A
91	472	30.8	295	31	US-10-170-205E-31797	Sequence 31797, A
92	472	30.8	295	31	US-10-170-205E-34133	Sequence 34133, A
93	472	30.8	295	31	US-10-170-205E-34134	Sequence 34134, A
94	472	30.8	295	33	US-10-348-546-4	Sequence 4, Appli
95	472	30.8	295	35	US-10-533-519-1480	Sequence 1480, Ap
96	472	30.8	295	37	US-10-769-507-4	Sequence 4, Appli
97	472	30.8	295	40	US-11-033-030-21	Sequence 21, Appl
98	472	30.8	295	50	US-60-568-073-1205	Sequence 1205, Ap
99	472	30.8	298	31	US-10-170-205E-17237	Sequence 17237, A
100	467	30.5	296	27	US-09-791-537-112794	Sequence 112794, A

ALIGNMENTS

RESULT 1

PCT-US01-41986-4

Query Match 100.0%; Score 1530; DB 1; Length 284;

Best Local Similarity 100.0%; Pred. No. 1.4e-154;

Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MAESEAEPTSPGFEFESKYFFHGVRLPPFCRGKMEIEIANFPVRPSDVWIVTPKSGTSL 60

1 MAESEAEPTSPGFEFESKYFFHGVRLPPFCRGKMEIEIANFPVRPSDVWIVTPKSGTSL 60

61 LOEVVYLVSGADPDEIGLMMNIDQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120

61 LOEVVYLVSGADPDEIGLMMNIDQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120

121 LHNGDSKVIYMARNPKDLVSVYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFHVQ 180

121 LHNGDSKVIYMARNPKDLVSVYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFHVQ 180

191 EFWEHRMDSNVFLKYEDMHRDLVTWVEQLARFLGVSCDKAQLALETCHQLVDQCCNA 240

191 EFWEHRMDSNVFLKYEDMHRDLVTWVEQLARFLGVSCDKAQLALETCHQLVDQCCNA 240

241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTTFDFYL 284

241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTTFDFYL 284

GENERAL INFORMATION:

APPLICANT: Trustees of Boston University

TITLE OF INVENTION: A NEURAL SPECIFIC CYTOSOLIC

TITLE OF INVENTION: SULFOTRANSFERASE FOR DRUG SCREENING

FILE REFERENCE: 13594-007W01

CURRENT APPLICATION NUMBER: PCT/US01/41986

CURRENT FILING DATE: 2001-09-04

PRIOR APPLICATION NUMBER: 60/229,929

PRIOR FILING DATE: 2000-09-01

NUMBER OF SEQ ID NOS: 42

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 6

LENGTH: 284

TYPE: PRT

ORGANISM: Homo sapiens

PCT-US01-41986-6

Query Match 100.0%; Score 1530; DB 1; Length 284;

Best Local Similarity 100.0%; Pred. No. 1.4e-154;

Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MAESEAEPTSPGFEFESKYFFHGVRLPPFCRGKMEIEIANFPVRPSDVWIVTPKSGTSL 60

1 MAESEAEPTSPGFEFESKYFFHGVRLPPFCRGKMEIEIANFPVRPSDVWIVTPKSGTSL 60

61 LOEVVYLVSGADPDEIGLMMNIDQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120

61 LOEVVYLVSGADPDEIGLMMNIDQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120

121 LHNGDSKVIYMARNPKDLVSVYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFHVQ 180

121 LHNGDSKVIYMARNPKDLVSVYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFHVQ 180

191 EFWEHRMDSNVFLKYEDMHRDLVTWVEQLARFLGVSCDKAQLALETCHQLVDQCCNA 240

191 EFWEHRMDSNVFLKYEDMHRDLVTWVEQLARFLGVSCDKAQLALETCHQLVDQCCNA 240

241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTTFDFYL 284

241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTTFDFYL 284

GENERAL INFORMATION:

APPLICANT: Agensys, Inc.

RESULT 3

PCT-US02-11654-60

Query Match 100.0%; Score 1530; DB 1; Length 284;

Best Local Similarity 100.0%; Pred. No. 1.4e-154;

Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MAESEAEPTSPGFEFESKYFFHGVRLPPFCRGKMEIEIANFPVRPSDVWIVTPKSGTSL 60

1 MAESEAEPTSPGFEFESKYFFHGVRLPPFCRGKMEIEIANFPVRPSDVWIVTPKSGTSL 60

61 LOEVVYLVSGADPDEIGLMMNIDQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120

61 LOEVVYLVSGADPDEIGLMMNIDQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120

121 LHNGDSKVIYMARNPKDLVSVYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFHVQ 180

121 LHNGDSKVIYMARNPKDLVSVYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFHVQ 180

191 EFWEHRMDSNVFLKYEDMHRDLVTWVEQLARFLGVSCDKAQLALETCHQLVDQCCNA 240

191 EFWEHRMDSNVFLKYEDMHRDLVTWVEQLARFLGVSCDKAQLALETCHQLVDQCCNA 240

241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTTFDFYL 284

241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTTFDFYL 284

GENERAL INFORMATION:

APPLICANT: Agensys, Inc.

```
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins Useful
; FILE REFERENCE: 51158-20040.40
; CURRENT APPLICATION NUMBER: PCT/US02/11654
; CURRENT FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US 60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-11654-60

Query Match      100.0%; Score 1530; DB 1; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MAESEAETPTGGEFESKYFFHGVRLPPFCRGKMEIEIANFPVRPSDVWIVTPKSGTSL 60
Db      1 MAESEAETPTGGEFESKYFFHGVRLPPFCRGKMEIEIANFPVRPSDVWIVTPKSGTSL 60

Qy      61 LOEVVYLVSQADPDEIGLNMIDQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db      61 LOEVVYLVSQADPDEIGLNMIDQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120

Qy      121 LHNGDSKVIYMARNPKDLVSVYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180
Db      121 LHNGDSKVIYMARNPKDLVSVYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180

Qy      181 EFWHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240
Db      181 EFWHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240

Qy      241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTDFYL 284
Db      241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTDFYL 284
```

```
RESULT 4
PCT-US02-11654-81
; Sequence 81, Application PC/TUS0211654
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins Useful
; FILE REFERENCE: 51158-20040.40
; CURRENT APPLICATION NUMBER: PCT/US02/11654
; CURRENT FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US 60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 103
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-11654-103

Query Match      100.0%; Score 1530; DB 1; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MAESEAETPTGGEFESKYFFHGVRLPPFCRGKMEIEIANFPVRPSDVWIVTPKSGTSL 60
```

```
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 81
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-11654-81

Query Match      100.0%; Score 1530; DB 1; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MAESEAETPTGGEFESKYFFHGVRLPPFCRGKMEIEIANFPVRPSDVWIVTPKSGTSL 60
Db      1 MAESEAETPTGGEFESKYFFHGVRLPPFCRGKMEIEIANFPVRPSDVWIVTPKSGTSL 60

Qy      61 LOEVVYLVSQADPDEIGLNMIDQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db      61 LOEVVYLVSQADPDEIGLNMIDQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120

Qy      121 LHNGDSKVIYMARNPKDLVSVYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180
Db      121 LHNGDSKVIYMARNPKDLVSVYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180

Qy      181 EFWHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240
Db      181 EFWHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240

Qy      241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTDFYL 284
Db      241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTDFYL 284
```

```
RESULT 5
PCT-US02-11654-103
; Sequence 103, Application PC/TUS0211654
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins Useful
; FILE REFERENCE: 51158-20040.40
; CURRENT APPLICATION NUMBER: PCT/US02/11654
; CURRENT FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US 60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 103
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-11654-103

Query Match      100.0%; Score 1530; DB 1; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MAESEAETPTGGEFESKYFFHGVRLPPFCRGKMEIEIANFPVRPSDVWIVTPKSGTSL 60
```

```
Db 1 MAESEAETPTGFEFSGKFFHGVRLPPFCRGKMEIEANFPVRPSDVIWVTPKSGTSL 60
Qy 61 LOEVVYLVSQADPDEIGLNMIDQLVLEYVPPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LOEVVYLVSQADPDEIGLNMIDQLVLEYVPPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Qy 121 LHNGDSKVIYMARNPDKLVVSYQFHRSLRTMSYRGTFQFCRRFMDKLGYSWFEHVQ 180
Db 121 LHNGDSKVIYMARNPDKLVVSYQFHRSLRTMSYRGTFQFCRRFMDKLGYSWFEHVQ 180
Qy 181 EFWHRMDSNVLFLKYEDMHRDLVTWVEQLARFLVSCDKAQLAEALTEHCHQLVDQCCNA 240
Db 181 EFWHRMDSNVLFLKYEDMHRDLVTWVEQLARFLVSCDKAQLAEALTEHCHQLVDQCCNA 240
Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTTFDFYL 284
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTTFDFYL 284

RESULT 6
PCT-US02-11654-104
; Sequence 104, Application PC/TUS0211654
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins Useful
; TITLE OF INVENTION: in the Detection and Treatment of Various Cancers
; FILE REFERENCE: 51158-20040.40
; CURRENT FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US 60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 104
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-11654-104

Query Match 100.0%; Score 1530; DB 1; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MAESEAETPTGFEFSGKFFHGVRLPPFCRGKMEIEANFPVRPSDVIWVTPKSGTSL 60
Db 1 MAESEAETPTGFEFSGKFFHGVRLPPFCRGKMEIEANFPVRPSDVIWVTPKSGTSL 60
Qy 61 LOEVVYLVSQADPDEIGLNMIDQLVLEYVPPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LOEVVYLVSQADPDEIGLNMIDQLVLEYVPPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Qy 121 LHNGDSKVIYMARNPDKLVVSYQFHRSLRTMSYRGTFQFCRRFMDKLGYSWFEHVQ 180
Db 121 LHNGDSKVIYMARNPDKLVVSYQFHRSLRTMSYRGTFQFCRRFMDKLGYSWFEHVQ 180
Qy 181 EFWHRMDSNVLFLKYEDMHRDLVTWVEQLARFLVSCDKAQLAEALTEHCHQLVDQCCNA 240
Db 181 EFWHRMDSNVLFLKYEDMHRDLVTWVEQLARFLVSCDKAQLAEALTEHCHQLVDQCCNA 240
Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTTFDFYL 284
```

```
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTTFDFYL 284

RESULT 7
PCT-US02-11654-708
; Sequence 708, Application PC/TUS0211654
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins Useful
; TITLE OF INVENTION: in the Detection and Treatment of Various Cancers
; FILE REFERENCE: 51158-20040.40
; CURRENT FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US 60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 708
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-11654-708

Query Match 100.0%; Score 1530; DB 1; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MAESEAETPTGFEFSGKFFHGVRLPPFCRGKMEIEANFPVRPSDVIWVTPKSGTSL 60
Db 1 MAESEAETPTGFEFSGKFFHGVRLPPFCRGKMEIEANFPVRPSDVIWVTPKSGTSL 60
Qy 61 LOEVVYLVSQADPDEIGLNMIDQLVLEYVPPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LOEVVYLVSQADPDEIGLNMIDQLVLEYVPPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Qy 121 LHNGDSKVIYMARNPDKLVVSYQFHRSLRTMSYRGTFQFCRRFMDKLGYSWFEHVQ 180
Db 121 LHNGDSKVIYMARNPDKLVVSYQFHRSLRTMSYRGTFQFCRRFMDKLGYSWFEHVQ 180
Qy 181 EFWHRMDSNVLFLKYEDMHRDLVTWVEQLARFLVSCDKAQLAEALTEHCHQLVDQCCNA 240
Db 181 EFWHRMDSNVLFLKYEDMHRDLVTWVEQLARFLVSCDKAQLAEALTEHCHQLVDQCCNA 240
Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTTFDFYL 284
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTTFDFYL 284

RESULT 8
PCT-US04-00750-110
; Sequence 110, Application PC/TUS0400750
; GENERAL INFORMATION:
; APPLICANT: Karicheti, Venkateswarlu
; APPLICANT: Silos-Santiago, Inmaculada
; APPLICANT: Eliasof, Scott D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: UROLOGICAL DISORDERS USING 44390, 54181, 211, 5687, 884,
; TITLE OF INVENTION: 1405, 636, 4421, 5410, 30905, 2045, 16405, 18560, 2047,
; TITLE OF INVENTION: 33751, 52872, 14063, 20739, 32544, 44373, 51164,
; TITLE OF INVENTION: 53010, 16852, 1587, 22245, 2387, 52908, 69112, 14990,
```

; TITLE OF INVENTION: 18547, 115, 579, 15985, 15625, 760, 18603, 2395, 2554, 8675,
; TITLE OF INVENTION: 32720, 4809, 14303, 16816, 17827, 32620, 577, 619, 1423,
; TITLE OF INVENTION: 2158, 8263, 15402, 16209, 16386, 21165, 30911, 41897, 1643,
; TITLE OF INVENTION: 2543, 9626, 13231, 32409, 84260, 2882, 8203, 32678 OR
; FILE REFERENCE: MPI03-007
; CURRENT APPLICATION NUMBER: PCT/US04/00750
; PRIOR FILING DATE: 2004-01-14
; PRIOR APPLICATION NUMBER: US 60/440,318
; PRIOR FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: US 60/444,783
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: US 60/457,901
; PRIOR FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: US 60/468,775
; PRIOR FILING DATE: 2003-05-08
; PRIOR APPLICATION NUMBER: US 60/471,614
; PRIOR FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: US 60/478,742
; PRIOR FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: US 60/488,529
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/491,156
; PRIOR FILING DATE: 2003-07-30
; PRIOR APPLICATION NUMBER: US 60/499,594
; PRIOR FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US 60/506,332
; PRIOR FILING DATE: 2003-09-26
; NUMBER OF SEQ ID NOS: 136
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 110
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US04-00750-110

Query Match 100.0%; Score 1530; DB 1; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAESEAETPTSGEPESKYFFHGVRLPPFCRGKMEIEANFPVRPSDVMIVTPKSGTSL 60
Db 1 MAESEAETPTSGEPESKYFFHGVRLPPFCRGKMEIEANFPVRPSDVMIVTPKSGTSL 60

Qy 61 LOEVVYLVSQADPDEIGLNMIDEQLPVLEYQPGLDIIKELTSPLIKSHLPYRFLPSD 120
Db 61 LOEVVYLVSQADPDEIGLNMIDEQLPVLEYQPGLDIIKELTSPLIKSHLPYRFLPSD 120

Qy 121 LHNGDSKVIYMARNPKDLVSVYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180
Db 121 LHNGDSKVIYMARNPKDLVSVYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180

Qy 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLALETCHQLVDQCCNA 240
Db 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLALETCHQLVDQCCNA 240

Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL 284
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL 284

RESULT 9
PCT-US04-02851-2
; Sequence 2, Application PC/TUS0402851
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Silos-Santiago, Inmaculada
; APPLICANT: Karicheti, Venkateswarlu
; APPLICANT: Eliassof, Scott D.
; TITLE OF INVENTION: PAIN AND PAINFUL DISORDERS USING 16386, 15402, 21165, 1423,
; TITLE OF INVENTION: 636, 12303, 21425, 27410, 38554, 38555, 55063, 57145, 59914,
; TITLE OF INVENTION: 94921, 16852, 33260, 58573, 30911, 85913, 14303, 16816,

; TITLE OF INVENTION: 17827 OR 32620
; FILE REFERENCE: MPI03-012
; CURRENT APPLICATION NUMBER: PCT/US04/02851
; CURRENT FILING DATE: 2004-01-30
; PRIOR APPLICATION NUMBER: US 60/444,781
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: US 60/452,291
; PRIOR FILING DATE: 2003-03-05
; PRIOR APPLICATION NUMBER: US 60/454,540
; PRIOR FILING DATE: 2003-03-13
; PRIOR APPLICATION NUMBER: US 60/478,805
; PRIOR FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: US 60/491,048
; PRIOR FILING DATE: 2003-07-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo Sapiens
PCT-US04-02851-2

Query Match 100.0%; Score 1530; DB 1; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAESEAETPTSGEPESKYFFHGVRLPPFCRGKMEIEANFPVRPSDVMIVTPKSGTSL 60
Db 1 MAESEAETPTSGEPESKYFFHGVRLPPFCRGKMEIEANFPVRPSDVMIVTPKSGTSL 60

Qy 61 LOEVVYLVSQADPDEIGLNMIDEQLPVLEYQPGLDIIKELTSPLIKSHLPYRFLPSD 120
Db 61 LOEVVYLVSQADPDEIGLNMIDEQLPVLEYQPGLDIIKELTSPLIKSHLPYRFLPSD 120

Qy 121 LHNGDSKVIYMARNPKDLVSVYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180
Db 121 LHNGDSKVIYMARNPKDLVSVYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180

Qy 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLALETCHQLVDQCCNA 240
Db 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLALETCHQLVDQCCNA 240

Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL 284
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL 284

RESULT 10
US-09-791-537-17119
; Sequence 17119, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17119
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-17119

Query Match 100.0%; Score 1530; DB 27; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAESEAETPTSGEPESKYFFHGVRLPPFCRGKMEIEANFPVRPSDVMIVTPKSGTSL 60

```
Db 1 MAESEATPTSGEPESKYFPHGVRLPPFCRGKMEIEANFVRPSDVWIVTPKSGTSL 60
Qy 61 LOEVVYLVSQADPDEIGLWNIDEQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LOEVVYLVSQADPDEIGLWNIDEQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120
Qy 121 LHNGDSKVIYMARNP KDVLVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHEHVQ 180
Db 121 LHNGDSKVIYMARNP KDVLVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHEHVQ 180
Qy 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVS CDKAQLEALTEHCHQLVDQCCNA 240
Db 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVS CDKAQLEALTEHCHQLVDQCCNA 240
Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTDFYL 284
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTDFYL 284

RESULT 11
US-10-121-024B-60
; Sequence 60, Application US/10121024B
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; FILE REFERENCE: 51158-20040.00
; CURRENT APPLICATION NUMBER: US/10/121,024B
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-121-024B-60

Query Match 100.0%; Score 1530; DB 31; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAESEATPTSGEPESKYFPHGVRLPPFCRGKMEIEANFVRPSDVWIVTPKSGTSL 60
Db 1 MAESEATPTSGEPESKYFPHGVRLPPFCRGKMEIEANFVRPSDVWIVTPKSGTSL 60
Qy 61 LOEVVYLVSQADPDEIGLWNIDEQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LOEVVYLVSQADPDEIGLWNIDEQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120
Qy 121 LHNGDSKVIYMARNP KDVLVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHEHVQ 180
Db 121 LHNGDSKVIYMARNP KDVLVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHEHVQ 180
Qy 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVS CDKAQLEALTEHCHQLVDQCCNA 240
Db 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVS CDKAQLEALTEHCHQLVDQCCNA 240
Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTDFYL 284
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTDFYL 284
```

```
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTDFYL 284

RESULT 12
US-10-121-024B-81
; Sequence 81, Application US/10121024B
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; FILE REFERENCE: 51158-20040.00
; CURRENT APPLICATION NUMBER: US/10/121,024B
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 81
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-121-024B-81

Query Match 100.0%; Score 1530; DB 31; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAESEATPTSGEPESKYFPHGVRLPPFCRGKMEIEANFVRPSDVWIVTPKSGTSL 60
Db 1 MAESEATPTSGEPESKYFPHGVRLPPFCRGKMEIEANFVRPSDVWIVTPKSGTSL 60
Qy 61 LOEVVYLVSQADPDEIGLWNIDEQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LOEVVYLVSQADPDEIGLWNIDEQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120
Qy 121 LHNGDSKVIYMARNP KDVLVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHEHVQ 180
Db 121 LHNGDSKVIYMARNP KDVLVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHEHVQ 180
Qy 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVS CDKAQLEALTEHCHQLVDQCCNA 240
Db 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVS CDKAQLEALTEHCHQLVDQCCNA 240
Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTDFYL 284
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTDFYL 284

RESULT 13
US-10-121-024B-103
; Sequence 103, Application US/10121024B
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
```

```
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; TITLE OF INVENTION: Useful in the Detection and Treatment of Various Cancers
; FILE REFERENCE: 51158-20040.00
; CURRENT APPLICATION NUMBER: US/10/121,024B
; CURRENT FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 103
; LENGTH: 284
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-121-024B-103

Query Match      100.0%; Score 1530; DB 31; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MAESEATPSTPGFESKYFEFHGVRLLPPFCRGKMEETIANPVRPSDVWIVTPKSGTSL 60
Db      1 MAESEATPSTPGFESKYFEFHGVRLLPPFCRGKMEETIANPVRPSDVWIVTPKSGTSL 60

Qy      61 LOEVVYLVSQGADPEIGLNMNIDQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db      61 LOEVVYLVSQGADPEIGLNMNIDQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120

Qy      121 LHNGDSKVIYMARNP KD LVSYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFHVQ 180
Db      121 LHNGDSKVIYMARNP KD LVSYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFHVQ 180

Qy      181 EFWEHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240
Db      181 EFWEHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240

Qy      241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL 284
Db      241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL 284

RESULT 14
US-10-121-024B-104
; Sequence 104, Application US/10121024B
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Paris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; TITLE OF INVENTION: Useful in the Detection and Treatment of Various Cancers
; FILE REFERENCE: 51158-20040.00
; CURRENT APPLICATION NUMBER: US/10/121,024B
; CURRENT FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 104
; LENGTH: 284
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-121-024B-104

Query Match      100.0%; Score 1530; DB 31; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MAESEATPSTPGFESKYFEFHGVRLLPPFCRGKMEETIANPVRPSDVWIVTPKSGTSL 60
Db      1 MAESEATPSTPGFESKYFEFHGVRLLPPFCRGKMEETIANPVRPSDVWIVTPKSGTSL 60

Qy      61 LOEVVYLVSQGADPEIGLNMNIDQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db      61 LOEVVYLVSQGADPEIGLNMNIDQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120

Qy      121 LHNGDSKVIYMARNP KD LVSYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFHVQ 180
Db      121 LHNGDSKVIYMARNP KD LVSYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFHVQ 180

Qy      181 EFWEHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240
Db      181 EFWEHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240

Qy      241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL 284
Db      241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL 284
```

```
; ORGANISM: Homo sapiens
US-10-121-024B-104

Query Match      100.0%; Score 1530; DB 31; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MAESEATPSTPGFESKYFEFHGVRLLPPFCRGKMEETIANPVRPSDVWIVTPKSGTSL 60
Db      1 MAESEATPSTPGFESKYFEFHGVRLLPPFCRGKMEETIANPVRPSDVWIVTPKSGTSL 60

Qy      61 LOEVVYLVSQGADPEIGLNMNIDQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db      61 LOEVVYLVSQGADPEIGLNMNIDQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120

Qy      121 LHNGDSKVIYMARNP KD LVSYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFHVQ 180
Db      121 LHNGDSKVIYMARNP KD LVSYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFHVQ 180

Qy      181 EFWEHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240
Db      181 EFWEHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240

Qy      241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL 284
Db      241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL 284

RESULT 15
US-10-121-024B-708
; Sequence 708, Application US/10121024B
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Paris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; TITLE OF INVENTION: Useful in the Detection and Treatment of Various Cancers
; FILE REFERENCE: 51158-20040.00
; CURRENT APPLICATION NUMBER: US/10/121,024B
; CURRENT FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 708
; LENGTH: 284
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-121-024B-708

Query Match      100.0%; Score 1530; DB 31; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MAESEATPSTPGFESKYFEFHGVRLLPPFCRGKMEETIANPVRPSDVWIVTPKSGTSL 60
Db      1 MAESEATPSTPGFESKYFEFHGVRLLPPFCRGKMEETIANPVRPSDVWIVTPKSGTSL 60

Qy      61 LOEVVYLVSQGADPEIGLNMNIDQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db      61 LOEVVYLVSQGADPEIGLNMNIDQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120

Qy      121 LHNGDSKVIYMARNP KD LVSYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFHVQ 180
Db      121 LHNGDSKVIYMARNP KD LVSYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFHVQ 180
```



```
Db 121 LHNGDSKVIYARNPKDLVSYQFHRSLRTWTSYAGTQFECRRPMNDKLGSGWFEHVQ 180
Qy 181 EFWEHRMDSNVLFKYEDMHRDLVTWVQDLARFLGVSCKQAQLEALTEHCHQLVDQCCNA 240
Db 181 EFWEHRMDSNVLFKYEDMHRDLVTWVQDLARFLGVSCKQAQLEALTEHCHQLVDQCCNA 240
Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284
```

RESULT 16

```
US-10-121-024C-60
; Sequence 60, Application US/10121024C
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Paris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins Useful
; TITLE OF INVENTION: in the Detection and Treatment of Various Cancers
; FILE REFERENCE: 51158-20040.00
; CURRENT APPLICATION NUMBER: US/10/121,024C
; CURRENT FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-121-024C-60
```

```
Query Match 100.0%; Score 1530; DB 31; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 MAESEATPSTPGFESKYFEGHGVRLPPFCRGKMEELANFPVRPSDVWIVTPKSGTSL 60
Db 1 MAESEATPSTPGFESKYFEGHGVRLPPFCRGKMEELANFPVRPSDVWIVTPKSGTSL 60
Qy 61 LQEVVYLVSQGADPDEIGLWNIDEQLPVLEYQPGGLDIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LQEVVYLVSQGADPDEIGLWNIDEQLPVLEYQPGGLDIKELTSPRLIKSHLPYRFLPSD 120
Qy 121 LHNGDSKVIYARNPKDLVSYQFHRSLRTWTSYAGTQFECRRPMNDKLGSGWFEHVQ 180
Db 121 LHNGDSKVIYARNPKDLVSYQFHRSLRTWTSYAGTQFECRRPMNDKLGSGWFEHVQ 180
Qy 181 EFWEHRMDSNVLFKYEDMHRDLVTWVQDLARFLGVSCKQAQLEALTEHCHQLVDQCCNA 240
Db 181 EFWEHRMDSNVLFKYEDMHRDLVTWVQDLARFLGVSCKQAQLEALTEHCHQLVDQCCNA 240
Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284
```

RESULT 17

```
US-10-121-024C-81
; Sequence 81, Application US/10121024C
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Aya
```

```
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Paris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins Useful
; TITLE OF INVENTION: in the Detection and Treatment of Various Cancers
; FILE REFERENCE: 51158-20040.00
; CURRENT APPLICATION NUMBER: US/10/121,024C
; CURRENT FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 81
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-121-024C-81
```

```
Query Match 100.0%; Score 1530; DB 31; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MAESEATPSTPGFESKYFEGHGVRLPPFCRGKMEELANFPVRPSDVWIVTPKSGTSL 60
Db 1 MAESEATPSTPGFESKYFEGHGVRLPPFCRGKMEELANFPVRPSDVWIVTPKSGTSL 60
Qy 61 LQEVVYLVSQGADPDEIGLWNIDEQLPVLEYQPGGLDIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LQEVVYLVSQGADPDEIGLWNIDEQLPVLEYQPGGLDIKELTSPRLIKSHLPYRFLPSD 120
Qy 121 LHNGDSKVIYARNPKDLVSYQFHRSLRTWTSYAGTQFECRRPMNDKLGSGWFEHVQ 180
Db 121 LHNGDSKVIYARNPKDLVSYQFHRSLRTWTSYAGTQFECRRPMNDKLGSGWFEHVQ 180
Qy 181 EFWEHRMDSNVLFKYEDMHRDLVTWVQDLARFLGVSCKQAQLEALTEHCHQLVDQCCNA 240
Db 181 EFWEHRMDSNVLFKYEDMHRDLVTWVQDLARFLGVSCKQAQLEALTEHCHQLVDQCCNA 240
Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284
```

RESULT 18

```
US-10-121-024C-103
; Sequence 103, Application US/10121024C
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Paris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins Useful
; TITLE OF INVENTION: in the Detection and Treatment of Various Cancers
; FILE REFERENCE: 51158-20040.00
; CURRENT APPLICATION NUMBER: US/10/121,024C
; CURRENT FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
```

		; PRIOR FILING DATE: 2001-04-25			
		; NUMBER OF SEQ ID NOS: 708			
		; SOFTWARE: FastSeq for Windows Version 4.0			
		; SEQ ID NO 103			
		; LENGTH: 284			
		; TYPE: PRT			
		; ORGANISM: Homo sapiens			
		US-10-121-024C-103			
		Query Match 100.0%; Score 1530; DB 31; Length 284;			
		Best Local Similarity 100.0%; Pred. No. 1.4e-154;			
		Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	MAESAETPSTPGEFESKYFFHGVRLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL	60		
Db	1	MAESAETPSTPGEFESKYFFHGVRLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL	60		
Qy	61	LQEVVYLVSQGADPDEIGLNMIDQLPVLEYQPQGLDIIKELTSPLRIKSHLPYRFLPSD	120		
Db	61	LQEVVYLVSQGADPDEIGLNMIDQLPVLEYQPQGLDIIKELTSPLRIKSHLPYRFLPSD	120		
Qy	121	LHNGDSKVIYARNPKDLVVSYYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVVQ	180		
Db	121	LHNGDSKVIYARNPKDLVVSYYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVVQ	180		
Qy	181	EFWEHRMDSNVLFLKYEDMHRDLVTMVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA	240		
Db	181	EFWEHRMDSNVLFLKYEDMHRDLVTMVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA	240		
Qy	241	EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL	284		
Db	241	EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL	284		
RESULT 20					
US-10-121-024C-708					
; Sequence 708, Application US/10121024C					
; GENERAL INFORMATION:					
; APPLICANT: Jakobovits, Aya					
; APPLICANT: Challita-Bid, Pia M.					
; APPLICANT: Faris, Mary					
; APPLICANT: Ge, Wangmao					
; APPLICANT: Hubert, Rene S.					
; APPLICANT: Morrison, Karen Jane Meyrick					
; APPLICANT: Morrison, Robert Kendall					
; APPLICANT: Raitano, Arthur B.					
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins Useful					
; FILE REFERENCE: 51158-20040.00					
; CURRENT APPLICATION NUMBER: US/10/121.024C					
; CURRENT FILING DATE: 2002-04-10					
; PRIOR APPLICATION NUMBER: US60/283,112					
; PRIOR FILING DATE: 2001-04-10					
; PRIOR APPLICATION NUMBER: US60/282,739					
; PRIOR FILING DATE: 2001-04-10					
; PRIOR APPLICATION NUMBER: US60/286,630					
; PRIOR FILING DATE: 2001-04-25					
; NUMBER OF SEQ ID NOS: 708					
; SOFTWARE: FastSeq for Windows Version 4.0					
; SEQ ID NO 708					
; LENGTH: 284					
; TYPE: PRT					
; ORGANISM: Homo sapiens					
US-10-121-024C-708					
		Query Match 100.0%; Score 1530; DB 31; Length 284;			
		Best Local Similarity 100.0%; Pred. No. 1.4e-154;			
		Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	MAESAETPSTPGEFESKYFFHGVRLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL	60		
Db	1	MAESAETPSTPGEFESKYFFHGVRLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL	60		
Qy	61	LQEVVYLVSQGADPDEIGLNMIDQLPVLEYQPQGLDIIKELTSPLRIKSHLPYRFLPSD	120		
Db	61	LQEVVYLVSQGADPDEIGLNMIDQLPVLEYQPQGLDIIKELTSPLRIKSHLPYRFLPSD	120		
Qy	121	LHNGDSKVIYARNPKDLVVSYYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVVQ	180		
Db	121	LHNGDSKVIYARNPKDLVVSYYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVVQ	180		
Qy	181	EFWEHRMDSNVLFLKYEDMHRDLVTMVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA	240		
Db	181	EFWEHRMDSNVLFLKYEDMHRDLVTMVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA	240		
Qy	241	EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL	284		
Db	241	EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL	284		
RESULT 19					
US-10-121-024C-104					
; Sequence 104, Application US/10121024C					
; GENERAL INFORMATION:					
; APPLICANT: Jakobovits, Aya					
; APPLICANT: Challita-Bid, Pia M.					
; APPLICANT: Faris, Mary					
; APPLICANT: Ge, Wangmao					
; APPLICANT: Hubert, Rene S.					
; APPLICANT: Morrison, Karen Jane Meyrick					
; APPLICANT: Morrison, Robert Kendall					
; APPLICANT: Raitano, Arthur B.					
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins Useful					
; FILE REFERENCE: 51158-20040.00					
; CURRENT APPLICATION NUMBER: US/10/121.024C					
; CURRENT FILING DATE: 2002-04-10					
; PRIOR APPLICATION NUMBER: US60/283,112					
; PRIOR FILING DATE: 2001-04-10					
; PRIOR APPLICATION NUMBER: US60/282,739					
; PRIOR FILING DATE: 2001-04-10					
; PRIOR APPLICATION NUMBER: US60/286,630					
; PRIOR FILING DATE: 2001-04-25					
; NUMBER OF SEQ ID NOS: 708					
; SOFTWARE: FastSeq for Windows Version 4.0					
; SEQ ID NO 104					
; LENGTH: 284					
; TYPE: PRT					
; ORGANISM: Homo sapiens					
US-10-121-024C-104					
		Query Match 100.0%; Score 1530; DB 31; Length 284;			
		Best Local Similarity 100.0%; Pred. No. 1.4e-154;			
		Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	MAESAETPSTPGEFESKYFFHGVRLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL	60		
Db	1	MAESAETPSTPGEFESKYFFHGVRLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL	60		

RESULT 21

US-10-427-631-7
; Sequence 7, Application US/10427631
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION; TANG, Y. TOM;
; APPLICANT: CORLEY, Neil C.; GUEGLER, Karl J.;
; APPLICANT: BAUGHN, Mariah R.; LAL, Preeti G.;
; APPLICANT: YUE, Henry; HILLMAN, Jennifer L.;
; APPLICANT: AZIMZAL, Yalda
; TITLE OF INVENTION: HUMAN TRANSFERASE PROTEINS
; FILE REFERENCE: PF-0592-1 DIV
; CURRENT APPLICATION NUMBER: US/10/427,631
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: US 09/786,240
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: PCT/US99/20989
; PRIOR FILING DATE: 1999-09-09
; PRIOR APPLICATION NUMBER: US 60/172,220
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: US 60/155,248
; PRIOR FILING DATE: 1998-11-04
; PRIOR APPLICATION NUMBER: US 60/133,642
; PRIOR FILING DATE: 1999-05-11
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PERL Program
; SEQ ID NO 7
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 1784742CD1
US-10-427-631-7

Query Match 100.0%; Score 1530; DB 34; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAESEAETPTGPEFSKYFPHGVRLPPFCRGKMEIANFPVRPSDVWVITYPKSGTSL 60
Db 1 MAESEAETPTGPEFSKYFPHGVRLPPFCRGKMEIANFPVRPSDVWVITYPKSGTSL 60

Qy 61 LQEVVYLSQAGADPEIGLWNIDQLPVLEYPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LQEVVYLSQAGADPEIGLWNIDQLPVLEYPQGLDIIKELTSPRLIKSHLPYRFLPSD 120

Qy 121 LHNGDSKVIYMARNPDLVSYIQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVQ 180
Db 121 LHNGDSKVIYMARNPDLVSYIQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVQ 180

Qy 181 EFWEHRMDSNVLFLKYEDMHRDLVTMVEQLARFLGVSCDKAQLTEHCHQLVDQCCNA 240
Db 181 EFWEHRMDSNVLFLKYEDMHRDLVTMVEQLARFLGVSCDKAQLTEHCHQLVDQCCNA 240

Qy 241 EALPVGRGRVGLWKDIFTVSNMEKFDLVYKQMGKCDLTTFDFYL 284
Db 241 EALPVGRGRVGLWKDIFTVSNMEKFDLVYKQMGKCDLTTFDFYL 284

RESULT 22

US-10-573-989-17
; Sequence 17, Application US/10573989
; GENERAL INFORMATION:
; APPLICANT: Evotec NeuroSciences GmbH
; APPLICANT: Von Der Kammer, Heinz
; APPLICANT: Pohner, Johannes
; TITLE OF INVENTION: DIAGNOSTIC AND THERAPEUTIC USE OF A SULFOTRANSFERASE
; FILE REFERENCE: 2335.0140000/SRL/KPQ
; CURRENT APPLICATION NUMBER: US/10/573,989
; CURRENT FILING DATE: 2006-03-30
; PRIOR APPLICATION NUMBER: PCT/EP2004/052353
; PRIOR FILING DATE: 2004-09-29

; PRIOR APPLICATION NUMBER: 60/506,775
; PRIOR FILING DATE: 2003-09-30
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-573-989-17

Query Match 100.0%; Score 1530; DB 35; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAESEAETPTGPEFSKYFPHGVRLPPFCRGKMEIANFPVRPSDVWVITYPKSGTSL 60
Db 1 MAESEAETPTGPEFSKYFPHGVRLPPFCRGKMEIANFPVRPSDVWVITYPKSGTSL 60

Qy 61 LQEVVYLSQAGADPEIGLWNIDQLPVLEYPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LQEVVYLSQAGADPEIGLWNIDQLPVLEYPQGLDIIKELTSPRLIKSHLPYRFLPSD 120

Qy 121 LHNGDSKVIYMARNPDLVSYIQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVQ 180
Db 121 LHNGDSKVIYMARNPDLVSYIQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVQ 180

Qy 181 EFWEHRMDSNVLFLKYEDMHRDLVTMVEQLARFLGVSCDKAQLTEHCHQLVDQCCNA 240
Db 181 EFWEHRMDSNVLFLKYEDMHRDLVTMVEQLARFLGVSCDKAQLTEHCHQLVDQCCNA 240

Qy 241 EALPVGRGRVGLWKDIFTVSNMEKFDLVYKQMGKCDLTTFDFYL 284
Db 241 EALPVGRGRVGLWKDIFTVSNMEKFDLVYKQMGKCDLTTFDFYL 284

RESULT 23

US-10-757-262-110
; Sequence 110, Application US/10757262
; GENERAL INFORMATION:
; APPLICANT: Karicheti, Venkateswarlu
; APPLICANT: Silos-Santiago, Immaculada
; APPLICANT: Eliasof, Scott D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: UROLOGICAL DISORDERS USING 44390, 54181, 211, 5687, 884,
; TITLE OF INVENTION: 1405, 636, 4421, 5410, 30905, 2045, 16405, 18580, 2047,
; TITLE OF INVENTION: 33751, 52872, 14063, 20739, 32544, 43239, 44373, 51164,
; TITLE OF INVENTION: 53010, 16852, 1587, 2207, 22245, 2387, 52908, 69112, 14990,
; TITLE OF INVENTION: 18547, 115, 579, 15985, 15625, 760, 18603, 2395, 2554, 8675,
; TITLE OF INVENTION: 32720, 4809, 14303, 16816, 17827, 32620, 577, 619, 1423,
; TITLE OF INVENTION: 2158, 8263, 15402, 16209, 16386, 21165, 30911, 41897, 1643,
; TITLE OF INVENTION: 2543, 9626, 13231, 32409, 84260, 2882, 8203, 32678 OR
; FILE REFERENCE: MPI03-007PIRNONMIN
; CURRENT APPLICATION NUMBER: US/10/757,262
; CURRENT FILING DATE: 2004-01-14
; PRIOR APPLICATION NUMBER: US 60/440,318
; PRIOR FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: US 60/444,783
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: US 60/457,901
; PRIOR FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: US 60/468,775
; PRIOR FILING DATE: 2003-05-08
; PRIOR APPLICATION NUMBER: US 60/471,614
; PRIOR FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: US 60/478,742
; PRIOR FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: US 60/488,529
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/491,156
; PRIOR FILING DATE: 2003-07-30
; PRIOR APPLICATION NUMBER: US 60/499,594
; PRIOR FILING DATE: 2003-09-02

;; PRIOR APPLICATION NUMBER: US 60/506,332
;; PRIOR FILING DATE: 2003-09-26
;; NUMBER OF SEQ ID NOS: 136
;; SOFTWARE: FASTSEQ for Windows Version 4.0
;; SEQ ID NO 110
;; LENGTH: 284
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-757-262-110

Query Match 100.0%; Score 1530; DB 37; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAESEAETPTGPEFESKYFEHGVRLPPFCRGKMEIEANFPVRPSDVMIVTPKSGTSL 60
DB 1 MAESEAETPTGPEFESKYFEHGVRLPPFCRGKMEIEANFPVRPSDVMIVTPKSGTSL 60

QY 61 LOEVVYLVVSQADPDEIGLNMNIDEOLPVLEYPOQGLDIIKELTSRLIKSHLPYRFLPSD 120
DB 61 LOEVVYLVVSQADPDEIGLNMNIDEOLPVLEYPOQGLDIIKELTSRLIKSHLPYRFLPSD 120

QY 121 LHNGDSKVIYMARNPDLVVSYYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180
DB 121 LHNGDSKVIYMARNPDLVVSYYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180

QY 181 EFWEHRMDSNVFLKYEDMHRDLVTMVEQLARFLGVSCDKAQLALETCHQLVDQCCNA 240
DB 181 EFWEHRMDSNVFLKYEDMHRDLVTMVEQLARFLGVSCDKAQLALETCHQLVDQCCNA 240

QY 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284
DB 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284

RESULT 24
US-10-768-158-2

;; Sequence 2, Application US/10768158
;; GENERAL INFORMATION:
;; APPLICANT: Millennium Pharmaceuticals, Inc.
;; APPLICANT: Silos-Santiago, Immaculada
;; APPLICANT: Karicheti, Venkateswarlu
;; APPLICANT: Eliasof, Scott D.
;; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
;; TITLE OF INVENTION: PAIN AND PAINFUL DISORDERS USING 16386, 15402, 21165, 1423,
;; TITLE OF INVENTION: 636, 12303, 21425, 27410, 38554, 38555, 55063, 57145, 59914,
;; TITLE OF INVENTION: 94921, 16852, 33260, 58573, 30911, 85913, 14303, 16816,
;; TITLE OF INVENTION: 17827 OR 32620
;; FILE REFERENCE: MP103-012P1RNMNIM
;; CURRENT APPLICATION NUMBER: US/10/768,158
;; PRIOR FILING DATE: 2004-01-30
;; PRIOR APPLICATION NUMBER: US 60/444,781
;; PRIOR FILING DATE: 2003-02-04
;; PRIOR APPLICATION NUMBER: US 60/452,291
;; PRIOR FILING DATE: 2003-03-05
;; PRIOR APPLICATION NUMBER: US 60/454,540
;; PRIOR FILING DATE: 2003-03-13
;; PRIOR APPLICATION NUMBER: US 60/478,805
;; PRIOR FILING DATE: 2003-06-16
;; PRIOR APPLICATION NUMBER: US 60/491,048
;; PRIOR FILING DATE: 2003-07-30
;; NUMBER OF SEQ ID NOS: 46
;; SOFTWARE: FASTSEQ for Windows Version 4.0
;; SEQ ID NO 2
;; LENGTH: 284
;; TYPE: PRT
;; ORGANISM: Homo Sapiens
US-10-768-158-2

Query Match 100.0%; Score 1530; DB 37; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAESEAETPTGPEFESKYFEHGVRLPPFCRGKMEIEANFPVRPSDVMIVTPKSGTSL 60
DB 1 MAESEAETPTGPEFESKYFEHGVRLPPFCRGKMEIEANFPVRPSDVMIVTPKSGTSL 60

QY 61 LOEVVYLVVSQADPDEIGLNMNIDEOLPVLEYPOQGLDIIKELTSRLIKSHLPYRFLPSD 120
DB 61 LOEVVYLVVSQADPDEIGLNMNIDEOLPVLEYPOQGLDIIKELTSRLIKSHLPYRFLPSD 120

QY 121 LHNGDSKVIYMARNPDLVVSYYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180
DB 121 LHNGDSKVIYMARNPDLVVSYYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180

QY 181 EFWEHRMDSNVFLKYEDMHRDLVTMVEQLARFLGVSCDKAQLALETCHQLVDQCCNA 240
DB 181 EFWEHRMDSNVFLKYEDMHRDLVTMVEQLARFLGVSCDKAQLALETCHQLVDQCCNA 240

QY 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284
DB 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284

RESULT 25
US-10-932-348-408

;; Sequence 408, Application US/10932348
;; GENERAL INFORMATION:
;; APPLICANT: CARGILL, Michele
;; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
;; TITLE OF INVENTION: ENCODING HUMAN DRUG-METABOLIZING ENZYMES, METHODS OF
;; TITLE OF INVENTION: DETECTION AND USES THEREOF
;; FILE REFERENCE: CL001481
;; CURRENT APPLICATION NUMBER: US/10/932,348
;; CURRENT FILING DATE: 2004-09-02
;; NUMBER OF SEQ ID NOS: 30431
;; SOFTWARE: FASTSEQ for Windows Version 4.0
;; SEQ ID NO 408
;; LENGTH: 284
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-932-348-408

Query Match 100.0%; Score 1530; DB 39; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAESEAETPTGPEFESKYFEHGVRLPPFCRGKMEIEANFPVRPSDVMIVTPKSGTSL 60
DB 1 MAESEAETPTGPEFESKYFEHGVRLPPFCRGKMEIEANFPVRPSDVMIVTPKSGTSL 60

QY 61 LOEVVYLVVSQADPDEIGLNMNIDEOLPVLEYPOQGLDIIKELTSRLIKSHLPYRFLPSD 120
DB 61 LOEVVYLVVSQADPDEIGLNMNIDEOLPVLEYPOQGLDIIKELTSRLIKSHLPYRFLPSD 120

QY 121 LHNGDSKVIYMARNPDLVVSYYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180
DB 121 LHNGDSKVIYMARNPDLVVSYYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180

QY 181 EFWEHRMDSNVFLKYEDMHRDLVTMVEQLARFLGVSCDKAQLALETCHQLVDQCCNA 240
DB 181 EFWEHRMDSNVFLKYEDMHRDLVTMVEQLARFLGVSCDKAQLALETCHQLVDQCCNA 240

QY 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284
DB 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284

RESULT 26
US-10-990-328-7686

;; Sequence 7686, Application US/10990328
;; GENERAL INFORMATION:
;; APPLICANT: CARGILL, Michele
;; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
;; TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
;; TITLE OF INVENTION: USES THEREOF

```
; FILE REFERENCE: CL001495
; CURRENT APPLICATION NUMBER: US/10/990,328
; CURRENT FILING DATE: 2004-11-17
; NUMBER OF SEQ ID NOS: 558824
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7686
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-990-328-7686

Query Match      100.0%; Score 1530; DB 39; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MAESEATPTPGEFESKYFEFHGVRLLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60
Db      1 MAESEATPTPGEFESKYFEFHGVRLLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60

Qy      61 LQEVVYLVSQADPDEIGLNNIDEQLPVLEYPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db      61 LQEVVYLVSQADPDEIGLNNIDEQLPVLEYPQGLDIIKELTSPRLIKSHLPYRFLPSD 120

Qy      121 LHNGDSKVIYARNPKDLVSYQFHRSLRTMSYRGTFQFCRRPMNDKLGYSWFEHVQ 180
Db      121 LHNGDSKVIYARNPKDLVSYQFHRSLRTMSYRGTFQFCRRPMNDKLGYSWFEHVQ 180

Qy      181 EFWEHRMDSNVFLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLALETCHQLVDQCCNA 240
Db      181 EFWEHRMDSNVFLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLALETCHQLVDQCCNA 240

Qy      241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTTFDFYL 284
Db      241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTTFDFYL 284

RESULT 27
US-10-990-328A-7686
; Sequence 7686, Application US/10990328A
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL001495
; CURRENT APPLICATION NUMBER: US/10/990,328A
; CURRENT FILING DATE: 2004-11-17
; NUMBER OF SEQ ID NOS: 558824
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7686
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-990-328A-7686

Query Match      100.0%; Score 1530; DB 39; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MAESEATPTPGEFESKYFEFHGVRLLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60
Db      1 MAESEATPTPGEFESKYFEFHGVRLLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60

Qy      61 LQEVVYLVSQADPDEIGLNNIDEQLPVLEYPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db      61 LQEVVYLVSQADPDEIGLNNIDEQLPVLEYPQGLDIIKELTSPRLIKSHLPYRFLPSD 120

Qy      121 LHNGDSKVIYARNPKDLVSYQFHRSLRTMSYRGTFQFCRRPMNDKLGYSWFEHVQ 180
Db      121 LHNGDSKVIYARNPKDLVSYQFHRSLRTMSYRGTFQFCRRPMNDKLGYSWFEHVQ 180

Qy      181 EFWEHRMDSNVFLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLALETCHQLVDQCCNA 240
Db      181 EFWEHRMDSNVFLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLALETCHQLVDQCCNA 240
```

```
Db      181 EFWEHRMDSNVFLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLALETCHQLVDQCCNA 240
Qy      241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTTFDFYL 284
Db      241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTTFDFYL 284

RESULT 28
US-11-154-298-60
; Sequence 60, Application US/11154298
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Paris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: METHODS OF INDUCING AN IMMUNE RESPONSE
; FILE REFERENCE: 51158-20040.01
; CURRENT APPLICATION NUMBER: US/11/154,298
; CURRENT FILING DATE: 2005-06-15
; PRIOR APPLICATION NUMBER: US10/121,024
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-154-298-60

Query Match      100.0%; Score 1530; DB 41; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MAESEATPTPGEFESKYFEFHGVRLLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60
Db      1 MAESEATPTPGEFESKYFEFHGVRLLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60

Qy      61 LQEVVYLVSQADPDEIGLNNIDEQLPVLEYPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db      61 LQEVVYLVSQADPDEIGLNNIDEQLPVLEYPQGLDIIKELTSPRLIKSHLPYRFLPSD 120

Qy      121 LHNGDSKVIYARNPKDLVSYQFHRSLRTMSYRGTFQFCRRPMNDKLGYSWFEHVQ 180
Db      121 LHNGDSKVIYARNPKDLVSYQFHRSLRTMSYRGTFQFCRRPMNDKLGYSWFEHVQ 180

Qy      181 EFWEHRMDSNVFLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLALETCHQLVDQCCNA 240
Db      181 EFWEHRMDSNVFLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLALETCHQLVDQCCNA 240

Qy      241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTTFDFYL 284
Db      241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTTFDFYL 284

RESULT 29
US-11-154-298-81
; Sequence 81, Application US/11154298
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Paris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
```

```
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: METHODS OF INDUCING AN IMMUNE RESPONSE
; FILE REFERENCE: 51158-20040.01
; CURRENT APPLICATION NUMBER: US/11/154,298
; CURRENT FILING DATE: 2005-06-15
; PRIOR APPLICATION NUMBER: US10/121,024
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 81
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-154-298-81

Query Match      100.0%; Score 1530; DB 41; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAESEAETPTGGEFESKYFEFHGVRLPPFCRGKMEIEANFVRPSDVWIVTPKSGTSL 60
DB 1 MAESEAETPTGGEFESKYFEFHGVRLPPFCRGKMEIEANFVRPSDVWIVTPKSGTSL 60
QY 61 LOEVVYLVSQGADPDEIGLMNIDEQLPVLEYPOGLDIIKELTSPLRIKSHLPYRFLPSD 120
DB 61 LOEVVYLVSQGADPDEIGLMNIDEQLPVLEYPOGLDIIKELTSPLRIKSHLPYRFLPSD 120
QY 121 LHNGDSKVIYMARNPCKDLVVSYYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180
DB 121 LHNGDSKVIYMARNPCKDLVVSYYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180
QY 181 EFWHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLTEALTECHQLVDQCCNA 240
DB 181 EFWHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLTEALTECHQLVDQCCNA 240
QY 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQXGKCDLTFDFYL 284
DB 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQXGKCDLTFDFYL 284

RESULT 31
US-11-154-298-104
; Sequence 104, Application US/11154298
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: METHODS OF INDUCING AN IMMUNE RESPONSE
; FILE REFERENCE: 51158-20040.01
; CURRENT APPLICATION NUMBER: US/11/154,298
; CURRENT FILING DATE: 2005-06-15
; PRIOR APPLICATION NUMBER: US10/121,024
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 104
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-154-298-104

Query Match      100.0%; Score 1530; DB 41; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAESEAETPTGGEFESKYFEFHGVRLPPFCRGKMEIEANFVRPSDVWIVTPKSGTSL 60
DB 1 MAESEAETPTGGEFESKYFEFHGVRLPPFCRGKMEIEANFVRPSDVWIVTPKSGTSL 60
QY 61 LOEVVYLVSQGADPDEIGLMNIDEQLPVLEYPOGLDIIKELTSPLRIKSHLPYRFLPSD 120
DB 61 LOEVVYLVSQGADPDEIGLMNIDEQLPVLEYPOGLDIIKELTSPLRIKSHLPYRFLPSD 120
QY 121 LHNGDSKVIYMARNPCKDLVVSYYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180
DB 121 LHNGDSKVIYMARNPCKDLVVSYYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180
QY 181 EFWHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLTEALTECHQLVDQCCNA 240
DB 181 EFWHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLTEALTECHQLVDQCCNA 240
QY 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQXGKCDLTFDFYL 284
DB 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQXGKCDLTFDFYL 284

RESULT 30
US-11-154-298-103
; Sequence 103, Application US/11154298
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: METHODS OF INDUCING AN IMMUNE RESPONSE
; FILE REFERENCE: 51158-20040.01
; CURRENT APPLICATION NUMBER: US/11/154,298
; CURRENT FILING DATE: 2005-06-15
; PRIOR APPLICATION NUMBER: US10/121,024
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
```

```
Db 61 LQEVVYLVSGQADPDEIGLNMIDQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Qy 121 LHNGDSKVIYARNPKDLVSVYQFHRSLRTMSYRGTFQFCRRPMNDKLGYSWFHVQ 180
Db 121 LHNGDSKVIYARNPKDLVSVYQFHRSLRTMSYRGTFQFCRRPMNDKLGYSWFHVQ 180
Qy 181 EFWHRMDSNVFLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLAETHCHQLVDQCCNA 240
Db 181 EFWHRMDSNVFLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLAETHCHQLVDQCCNA 240
Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL 284
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL 284

RESULT 32
US-11-154-298-708
; Sequence 708, Application US/11154298
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: METHODS OF INDUCING AN IMMUNE RESPONSE
; FILE REFERENCE: 51158-20040.01
; CURRENT APPLICATION NUMBER: US/11/154,298
; CURRENT FILING DATE: 2005-06-15
; PRIOR APPLICATION NUMBER: US10/121,024
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 708
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-154-298-708

Query Match 100.0%; Score 1530; DB 41; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAESEAEPTSTPGFESKYFEGHGVRLPPFCRGKMEIEANFPVRPSDVWIVTPKSGTSL 60
Db 1 MAESEAEPTSTPGFESKYFEGHGVRLPPFCRGKMEIEANFPVRPSDVWIVTPKSGTSL 60
Qy 61 LQEVVYLVSGQADPDEIGLNMIDQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LQEVVYLVSGQADPDEIGLNMIDQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Qy 121 LHNGDSKVIYARNPKDLVSVYQFHRSLRTMSYRGTFQFCRRPMNDKLGYSWFHVQ 180
Db 121 LHNGDSKVIYARNPKDLVSVYQFHRSLRTMSYRGTFQFCRRPMNDKLGYSWFHVQ 180
Qy 181 EFWHRMDSNVFLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLAETHCHQLVDQCCNA 240
Db 181 EFWHRMDSNVFLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLAETHCHQLVDQCCNA 240
Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL 284
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL 284

US-11-154-298-708
```

```
RESULT 33
US-11-155-906-60
; Sequence 60, Application US/11155906
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Raitano, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: METHODS OF INDUCING AN IMMUNE RESPONSE
; FILE REFERENCE: 51158-20040.02
; CURRENT APPLICATION NUMBER: US/11/155,906
; CURRENT FILING DATE: 2005-06-17
; PRIOR APPLICATION NUMBER: US10/121,024
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-155-906-60

Query Match 100.0%; Score 1530; DB 41; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAESEAEPTSTPGFESKYFEGHGVRLPPFCRGKMEIEANFPVRPSDVWIVTPKSGTSL 60
Db 1 MAESEAEPTSTPGFESKYFEGHGVRLPPFCRGKMEIEANFPVRPSDVWIVTPKSGTSL 60
Qy 61 LQEVVYLVSGQADPDEIGLNMIDQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LQEVVYLVSGQADPDEIGLNMIDQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Qy 121 LHNGDSKVIYARNPKDLVSVYQFHRSLRTMSYRGTFQFCRRPMNDKLGYSWFHVQ 180
Db 121 LHNGDSKVIYARNPKDLVSVYQFHRSLRTMSYRGTFQFCRRPMNDKLGYSWFHVQ 180
Qy 181 EFWHRMDSNVFLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLAETHCHQLVDQCCNA 240
Db 181 EFWHRMDSNVFLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLAETHCHQLVDQCCNA 240
Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL 284
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL 284

US-11-155-906-81
; Sequence 81, Application US/11155906
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Raitano, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: METHODS OF INDUCING AN IMMUNE RESPONSE
; FILE REFERENCE: 51158-20040.02
; CURRENT APPLICATION NUMBER: US/11/155,906
; CURRENT FILING DATE: 2005-06-17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 81
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-155-906-81
```

RESULT 34

```
US-11-155-906-81
; Sequence 81, Application US/11155906
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Raitano, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: METHODS OF INDUCING AN IMMUNE RESPONSE
; FILE REFERENCE: 51158-20040.02
; CURRENT APPLICATION NUMBER: US/11/155,906
; CURRENT FILING DATE: 2005-06-17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 81
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-155-906-81
```



```
; PRIOR APPLICATION NUMBER: US10/121,024
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 81
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-155-906-81

Query Match      100.0%; Score 1530; DB 41; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MAESEAETPSTPGEFESKYFEFHGVRLLPPFCRGKMEIEIANFPVRPSDVWIVTPKSGTSL 60
Db      1 MAESEAETPSTPGEFESKYFEFHGVRLLPPFCRGKMEIEIANFPVRPSDVWIVTPKSGTSL 60

Qy      61 LQEVVYLVSQGADPDEIGLMNIDQLPVLEYPOQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db      61 LQEVVYLVSQGADPDEIGLMNIDQLPVLEYPOQGLDIIKELTSPRLIKSHLPYRFLPSD 120

Qy      121 LHNGDSKVIYMARNP KDVLVSYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFHVQ 180
Db      121 LHNGDSKVIYMARNP KDVLVSYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFHVQ 180

Qy      181 EFWEHRMDSNVLFKYEDMHRDLVTWVQLARFLGVSCDKAQLALETCHQLVDQCCNA 240
Db      181 EFWEHRMDSNVLFKYEDMHRDLVTWVQLARFLGVSCDKAQLALETCHQLVDQCCNA 240

Qy      241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284
Db      241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284

RESULT 36
US-11-155-906-104
; Sequence 104, Application US/11155906
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Farris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: METHODS OF INDUCING AN IMMUNE RESPONSE
; FILE REFERENCE: 51158-20040.02
; CURRENT APPLICATION NUMBER: US/11/155,906
; CURRENT FILING DATE: 2005-06-17
; PRIOR APPLICATION NUMBER: US10/121,024
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 104
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-155-906-104

Query Match      100.0%; Score 1530; DB 41; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MAESEAETPSTPGEFESKYFEFHGVRLLPPFCRGKMEIEIANFPVRPSDVWIVTPKSGTSL 60
Db      1 MAESEAETPSTPGEFESKYFEFHGVRLLPPFCRGKMEIEIANFPVRPSDVWIVTPKSGTSL 60

Qy      61 LQEVVYLVSQGADPDEIGLMNIDQLPVLEYPOQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db      61 LQEVVYLVSQGADPDEIGLMNIDQLPVLEYPOQGLDIIKELTSPRLIKSHLPYRFLPSD 120

Qy      121 LHNGDSKVIYMARNP KDVLVSYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFHVQ 180
Db      121 LHNGDSKVIYMARNP KDVLVSYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFHVQ 180

Qy      181 EFWEHRMDSNVLFKYEDMHRDLVTWVQLARFLGVSCDKAQLALETCHQLVDQCCNA 240
Db      181 EFWEHRMDSNVLFKYEDMHRDLVTWVQLARFLGVSCDKAQLALETCHQLVDQCCNA 240

Qy      241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284
Db      241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284

RESULT 35
US-11-155-906-103
; Sequence 103, Application US/11155906
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Farris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: METHODS OF INDUCING AN IMMUNE RESPONSE
; FILE REFERENCE: 51158-20040.02
; CURRENT APPLICATION NUMBER: US/11/155,906
; CURRENT FILING DATE: 2005-06-17
; PRIOR APPLICATION NUMBER: US10/121,024
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 103
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-155-906-103
```

```

QY 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240
Db 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240

QY 241 EALPVGGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFFDYL 284
Db 241 EALPVGGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFFDYL 284

RESULT 37
US-11-155-906-708
; Sequence 708, Application US/11155906
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: METHODS OF INDUCING AN IMMUNE RESPONSE
; FILE REFERENCE: 51158-20040.02
; CURRENT APPLICATION NUMBER: US/11/155,906
; CURRENT FILING DATE: 2005-06-17
; PRIOR APPLICATION NUMBER: US10/121,024
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 708
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-155-906-708

Query Match 100.0%; Score 1530; DB 41; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAESEATPSTPGFESKYFEFHGVRLLPPFCRGKMEIANPVRPSDVWIVTPKSGTSL 60
Db 1 MAESEATPSTPGFESKYFEFHGVRLLPPFCRGKMEIANPVRPSDVWIVTPKSGTSL 60

QY 61 LOEVVYLVSGQADPDEIGLMNIDQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LOEVVYLVSGQADPDEIGLMNIDQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120

QY 121 LHNGDSKVIYMARNP KDVLVSYQFHRSLRTMSYRGTFQECRRFMDKLGYSWFEHVQ 180
Db 121 LHNGDSKVIYMARNP KDVLVSYQFHRSLRTMSYRGTFQECRRFMDKLGYSWFEHVQ 180

QY 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240
Db 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240

QY 241 EALPVGGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFFDYL 284
Db 241 EALPVGGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFFDYL 284

RESULT 38
US-11-156-231-60
; Sequence 60, Application US/11156231
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: METHODS OF INDUCING AN IMMUNE RESPONSE
; FILE REFERENCE: 51158-20040.03
; CURRENT APPLICATION NUMBER: US/11/156,231
; CURRENT FILING DATE: 2005-06-17
; PRIOR APPLICATION NUMBER: US10/121,024
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630

```

```

; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: METHODS OF INDUCING AN IMMUNE RESPONSE
; FILE REFERENCE: 51158-20040.03
; CURRENT APPLICATION NUMBER: US/11/156,231
; CURRENT FILING DATE: 2005-06-17
; PRIOR APPLICATION NUMBER: US10/121,024
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-156-231-60

Query Match 100.0%; Score 1530; DB 41; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAESEATPSTPGFESKYFEFHGVRLLPPFCRGKMEIANPVRPSDVWIVTPKSGTSL 60
Db 1 MAESEATPSTPGFESKYFEFHGVRLLPPFCRGKMEIANPVRPSDVWIVTPKSGTSL 60

QY 61 LOEVVYLVSGQADPDEIGLMNIDQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LOEVVYLVSGQADPDEIGLMNIDQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120

QY 121 LHNGDSKVIYMARNP KDVLVSYQFHRSLRTMSYRGTFQECRRFMDKLGYSWFEHVQ 180
Db 121 LHNGDSKVIYMARNP KDVLVSYQFHRSLRTMSYRGTFQECRRFMDKLGYSWFEHVQ 180

QY 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240
Db 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240

QY 241 EALPVGGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFFDYL 284
Db 241 EALPVGGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFFDYL 284

RESULT 39
US-11-156-231-81
; Sequence 81, Application US/11156231
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: METHODS OF INDUCING AN IMMUNE RESPONSE
; FILE REFERENCE: 51158-20040.03
; CURRENT APPLICATION NUMBER: US/11/156,231
; CURRENT FILING DATE: 2005-06-17
; PRIOR APPLICATION NUMBER: US10/121,024
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630

```

```
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 81
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-156-231-81

Query Match      100.0%; Score 1530; DB 41; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAESEAETPSTGPEFESKYFEFHGVLRLPPFCRGKMEETANFPVRPSDVMIVTPKSGTSL 60
DB 1 MAESEAETPSTGPEFESKYFEFHGVLRLPPFCRGKMEETANFPVRPSDVMIVTPKSGTSL 60

QY 61 LQEVVYLVSQGADPDEIGLNMNIDEQLPVLEYQPGLDIIKELTSRLIKSHLPYRFLPSD 120
DB 61 LQEVVYLVSQGADPDEIGLNMNIDEQLPVLEYQPGLDIIKELTSRLIKSHLPYRFLPSD 120

QY 121 LHNGDSKVIYMARNP KD LVVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180
DB 121 LHNGDSKVIYMARNP KD LVVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180

QY 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240
DB 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240

QY 241 EALPVGGRGVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284
DB 241 EALPVGGRGVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284

RESULT 40
US-11-156-231-103
; Sequence 103, Application US/11156231
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Ava
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Paris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: METHODS OF INDUCING AN IMMUNE RESPONSE
; FILE REFERENCE: 51158-20040.03
; CURRENT APPLICATION NUMBER: US10/121,024
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 104
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-156-231-104

Query Match      100.0%; Score 1530; DB 41; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAESEAETPSTGPEFESKYFEFHGVLRLPPFCRGKMEETANFPVRPSDVMIVTPKSGTSL 60
DB 1 MAESEAETPSTGPEFESKYFEFHGVLRLPPFCRGKMEETANFPVRPSDVMIVTPKSGTSL 60

QY 61 LQEVVYLVSQGADPDEIGLNMNIDEQLPVLEYQPGLDIIKELTSRLIKSHLPYRFLPSD 120
DB 61 LQEVVYLVSQGADPDEIGLNMNIDEQLPVLEYQPGLDIIKELTSRLIKSHLPYRFLPSD 120

QY 121 LHNGDSKVIYMARNP KD LVVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180
DB 121 LHNGDSKVIYMARNP KD LVVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180

QY 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240
DB 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240

QY 241 EALPVGGRGVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284
DB 241 EALPVGGRGVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284

RESULT 40
US-11-156-231-103
; Sequence 103, Application US/11156231
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Ava
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Paris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: METHODS OF INDUCING AN IMMUNE RESPONSE
; FILE REFERENCE: 51158-20040.03
; CURRENT APPLICATION NUMBER: US10/121,024
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 103
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-156-231-103

Query Match      100.0%; Score 1530; DB 41; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAESEAETPSTGPEFESKYFEFHGVLRLPPFCRGKMEETANFPVRPSDVMIVTPKSGTSL 60
DB 1 MAESEAETPSTGPEFESKYFEFHGVLRLPPFCRGKMEETANFPVRPSDVMIVTPKSGTSL 60
```



```
Db      241  EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284
|||||
RESULT 45
US-60-499-964-408
; Sequence 408, Application US/60499964
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN DRUG-METABOLIZING ENZYMES, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001481
; CURRENT APPLICATION NUMBER: US/60/499,964
; CURRENT FILING DATE: 2003-09-04
; NUMBER OF SEQ ID NOS: 30431
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 408
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-499-964-408

Query Match      100.0%; Score 1530; DB 49; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  MAESEAEPTSPGGEFESKYFHHGVRLLPPFCRGKMEETANFVRPSDVWIVTPKSGTSL 60
|||||
Db      1  MAESEAEPTSPGGEFESKYFHHGVRLLPPFCRGKMEETANFVRPSDVWIVTPKSGTSL 60
|||||
Qy      61  LOEVVYLVSQGADPDEIGLNMNIDEQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
|||||
Db      61  LOEVVYLVSQGADPDEIGLNMNIDEQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
|||||
Qy      121  LHNGDSKVIYMARNPKDLVVSYYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFHVVQ 180
|||||
Db      121  LHNGDSKVIYMARNPKDLVVSYYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFHVVQ 180
|||||
Qy      181  EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCKAQLAALTEHCHQLVDQCCNA 240
|||||
Db      181  EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCKAQLAALTEHCHQLVDQCCNA 240
|||||
Qy      241  EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284
|||||
Db      241  EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284
|||||

RESULT 46
US-10-170-205E-16293
; Sequence 16293, Application US/10170205E
; GENERAL INFORMATION:
; APPLICANT: ADAMS, Mark
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN
; TITLE OF INVENTION: CAPTURE AGENTS, AND USES THEREOF
; FILE REFERENCE: CL001381
; CURRENT APPLICATION NUMBER: US/10/170,205E
; CURRENT FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 40312
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 16293
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-170-205E-16293

Query Match      100.0%; Score 1530; DB 31; Length 311;
Best Local Similarity 100.0%; Pred. No. 1.6e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  MAESEAEPTSPGGEFESKYFHHGVRLLPPFCRGKMEETANFVRPSDVWIVTPKSGTSL 60
|||||
Db      28  MAESEAEPTSPGGEFESKYFHHGVRLLPPFCRGKMEETANFVRPSDVWIVTPKSGTSL 87
|||||

Db      241  EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284
|||||
Db      241  EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284
|||||

RESULT 47
US-60-245-241-166
; Sequence 166, Application US/60245241
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN DRUG TARGET ENZYMES,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN DRUG TARGET ENZYMES,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000888
; CURRENT APPLICATION NUMBER: US/60/245,241
; CURRENT FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 255
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 166
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Human
US-60-245-241-166

Query Match      100.0%; Score 1530; DB 47; Length 311;
Best Local Similarity 100.0%; Pred. No. 1.6e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  MAESEAEPTSPGGEFESKYFHHGVRLLPPFCRGKMEETANFVRPSDVWIVTPKSGTSL 60
|||||
Db      28  MAESEAEPTSPGGEFESKYFHHGVRLLPPFCRGKMEETANFVRPSDVWIVTPKSGTSL 87
|||||
Qy      61  LOEVVYLVSQGADPDEIGLNMNIDEQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
|||||
Db      88  LOEVVYLVSQGADPDEIGLNMNIDEQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 147
|||||
Qy      121  LHNGDSKVIYMARNPKDLVVSYYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFHVVQ 180
|||||
Db      148  LHNGDSKVIYMARNPKDLVVSYYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFHVVQ 207
|||||
Qy      181  EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCKAQLAALTEHCHQLVDQCCNA 240
|||||
Db      208  EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCKAQLAALTEHCHQLVDQCCNA 267
|||||
Qy      241  EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284
|||||
Db      268  EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 311
|||||

RESULT 48
US-60-258-016-46
; Sequence 46, Application US/60258016
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN DRUG TARGET ENZYMES,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN DRUG TARGET ENZYMES,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001037-PROV
; CURRENT APPLICATION NUMBER: US/60/258,016
; CURRENT FILING DATE: 2000-12-27
```

```
Qy      61  LOEVVYLVSQGADPDEIGLNMNIDEQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
|||||
Db      88  LOEVVYLVSQGADPDEIGLNMNIDEQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 147
|||||
Qy      121  LHNGDSKVIYMARNPKDLVVSYYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFHVVQ 180
|||||
Db      148  LHNGDSKVIYMARNPKDLVVSYYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFHVVQ 207
|||||
Qy      181  EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCKAQLAALTEHCHQLVDQCCNA 240
|||||
Db      208  EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCKAQLAALTEHCHQLVDQCCNA 267
|||||
Qy      241  EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284
|||||
Db      268  EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 311
|||||

RESULT 47
US-60-245-241-166
; Sequence 166, Application US/60245241
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN DRUG TARGET ENZYMES,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN DRUG TARGET ENZYMES,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000888
; CURRENT APPLICATION NUMBER: US/60/245,241
; CURRENT FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 255
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 166
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Human
US-60-245-241-166

Query Match      100.0%; Score 1530; DB 47; Length 311;
Best Local Similarity 100.0%; Pred. No. 1.6e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  MAESEAEPTSPGGEFESKYFHHGVRLLPPFCRGKMEETANFVRPSDVWIVTPKSGTSL 60
|||||
Db      28  MAESEAEPTSPGGEFESKYFHHGVRLLPPFCRGKMEETANFVRPSDVWIVTPKSGTSL 87
|||||
Qy      61  LOEVVYLVSQGADPDEIGLNMNIDEQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
|||||
Db      88  LOEVVYLVSQGADPDEIGLNMNIDEQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 147
|||||
Qy      121  LHNGDSKVIYMARNPKDLVVSYYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFHVVQ 180
|||||
Db      148  LHNGDSKVIYMARNPKDLVVSYYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFHVVQ 207
|||||
Qy      181  EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCKAQLAALTEHCHQLVDQCCNA 240
|||||
Db      208  EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCKAQLAALTEHCHQLVDQCCNA 267
|||||
Qy      241  EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284
|||||
Db      268  EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 311
|||||

RESULT 48
US-60-258-016-46
; Sequence 46, Application US/60258016
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN DRUG TARGET ENZYMES,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN DRUG TARGET ENZYMES,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001037-PROV
; CURRENT APPLICATION NUMBER: US/60/258,016
; CURRENT FILING DATE: 2000-12-27
```

; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Human
US-60-258-016-46

Query Match 100.0%; Score 1530; DB 47; Length 311;
Best Local Similarity 100.0%; Pred. No. 1.6e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAESEAEPTSTPGEFESKYFEFGVRLPPFCRGMKEEIANFPVRPSDVWIVTPKSGTSL 60
Db 28 MAESEAEPTSTPGEFESKYFEFGVRLPPFCRGMKEEIANFPVRPSDVWIVTPKSGTSL 87
Qy 61 LQEVVYLSQADPDEIGLMNIDEQLPVLEYPOQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 88 LQEVVYLSQADPDEIGLMNIDEQLPVLEYPOQGLDIIKELTSPRLIKSHLPYRFLPSD 147
Qy 121 LHNGDSKVIYARNPKDLVVSYYQFHRSLRTMSYRGTFQECRRPMNDKLGYSWFHVQ 180
Db 148 LHNGDSKVIYARNPKDLVVSYYQFHRSLRTMSYRGTFQECRRPMNDKLGYSWFHVQ 207
Qy 181 EFWEHRMDSNVLFLKYEDMHRDLVTWVEQLARFLGVSCKAQLALETCHQLVDQCCNA 240
Db 208 EFWEHRMDSNVLFLKYEDMHRDLVTWVEQLARFLGVSCKAQLALETCHQLVDQCCNA 267
Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTDFDYL 284
Db 268 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTDFDYL 311

RESULT 49

US-09-791-537-89260
; Sequence 89260, Application US/09791537
; GENERAL INFORMATION:

; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 89260
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-791-537-89260

Query Match 98.4%; Score 1506; DB 27; Length 284;
Best Local Similarity 97.9%; Pred. No. 5.3e-152;
Matches 278; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MAESEAEPTSTPGEFESKYFEFGVRLPPFCRGMKEEIANFPVRPSDVWIVTPKSGTSL 60
Db 1 MAESEAEPTSTPGEFESKYFEFGVRLPPFCRGMKEEIANFPVRPSDVWIVTPKSGTSL 60
Qy 61 LQEVVYLSQADPDEIGLMNIDEQLPVLEYPOQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LQEVVYLSQADPDEIGLMNIDEQLPVLEYPOQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Qy 121 LHNGDSKVIYARNPKDLVVSYYQFHRSLRTMSYRGTFQECRRPMNDKLGYSWFHVQ 180
Db 121 LHNGDSKVIYARNPKDLVVSYYQFHRSLRTMSYRGTFQECRRPMNDKLGYSWFHVQ 180
Qy 181 EFWEHRMDSNVLFLKYEDMHRDLVTWVEQLARFLGVSCKAQLALETCHQLVDQCCNA 240
Db 181 EFWEHRMDSNVLFLKYEDMHRDLVTWVEQLARFLGVSCKAQLALETCHQLVDQCCNA 240

Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTDFDYL 284
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTDFDYL 284

RESULT 50

US-10-205-331-46

; Sequence 46, Application US/10205331
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; APPLICANT: Lee, Kevin
; APPLICANT: Dixon, Alistair
; APPLICANT: Brooksbank, Robert
; APPLICANT: Pincock, Robert
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
; FILE REFERENCE: WL-A-018199
; CURRENT APPLICATION NUMBER: US/10/205,331
; CURRENT FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: GB 0118354.0
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 117
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 46
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Sulfotransferase-like protein
US-10-205-331-46

Query Match 98.4%; Score 1506; DB 32; Length 284;
Best Local Similarity 97.9%; Pred. No. 5.3e-152;
Matches 278; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MAESEAEPTSTPGEFESKYFEFGVRLPPFCRGMKEEIANFPVRPSDVWIVTPKSGTSL 60
Db 1 MAESEAEPTSTPGEFESKYFEFGVRLPPFCRGMKEEIANFPVRPSDVWIVTPKSGTSL 60
Qy 61 LQEVVYLSQADPDEIGLMNIDEQLPVLEYPOQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LQEVVYLSQADPDEIGLMNIDEQLPVLEYPOQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Qy 121 LHNGDSKVIYARNPKDLVVSYYQFHRSLRTMSYRGTFQECRRPMNDKLGYSWFHVQ 180
Db 121 LHNGDSKVIYARNPKDLVVSYYQFHRSLRTMSYRGTFQECRRPMNDKLGYSWFHVQ 180
Qy 181 EFWEHRMDSNVLFLKYEDMHRDLVTWVEQLARFLGVSCKAQLALETCHQLVDQCCNA 240
Db 181 EFWEHRMDSNVLFLKYEDMHRDLVTWVEQLARFLGVSCKAQLALETCHQLVDQCCNA 240
Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTDFDYL 284
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTDFDYL 284

RESULT 51

US-60-229-512-225

; Sequence 225, Application US/60229512
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PHASE II
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, NUCLEIC ACID MOLECULES ENCODING
; TITLE OF INVENTION: HUMAN DRUG-METABOLIZING PHASE II PROTEINS, AND USES THEREOF
; FILE REFERENCE: CU000770
; CURRENT APPLICATION NUMBER: US/60/229,512
; CURRENT FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 516
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 225
; LENGTH: 319
; TYPE: PRT
; ORGANISM: HUMAN
US-60-229-512-225

```
Query Match      98.2%; Score 1502.5; DB 47; Length 319;
Best Local Similarity 89.0%; Pred. No. 1.5e-151;
Matches 284; Conservative 0; Mismatches 0; Indels 35; Gaps 1;

QY 1 MAESAETPSTGPEFESKYFEHGVRLPPFCRGMKEEIANFPVRPSDVWIVTPKSGTSL 56
DB 1 MAESAETPSTGPEFESKYFEHGVRLPPFCRGMKEEIANFPVRPSDVWIVTPKSGRHH 60
QY 57 -----GTSLLQEVVYLVSOQADPDEIGLWNIDQ 85
DB 61 AGRDEEPAETGAADMGKQWVKLCPFLGTSLLQEVVYLVSOQADPDEIGLWNIDQ 120
QY 86 LPVLYPQGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPKDLVVSYYOF 145
DB 121 LPVLYPQGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPKDLVVSYYOF 180
QY 146 HRSRTMSYRGTFQFCRRFMNDKLGYSWFHGVQEFWEHRMDSNVFLFKYEDMHRDLVT 205
DB 181 HRSRTMSYRGTFQFCRRFMNDKLGYSWFHGVQEFWEHRMDSNVFLFKYEDMHRDLVT 240
QY 206 MVEQLARFLGVSCDKAQLAETHECHQVLDQCCNAEALPVGRVGLWKDIFTVSNNEKF 265
DB 241 MVEQLARFLGVSCDKAQLAETHECHQVLDQCCNAEALPVGRVGLWKDIFTVSNNEKF 300
QY 266 DLVYKQKMGKCDLTDFDYL 284
DB 301 DLVYKQKMGKCDLTDFDYL 319

RESULT 52
US-10-170-205E-27511
; Sequence 27511, Application US/10170205E
; GENERAL INFORMATION:
; APPLICANT: ADAMS, Mark
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN
; FILE OF INVENTION: CAPTURE AGENTS, AND USES THEREOF
; FILE REFERENCE: CL001381
; CURRENT APPLICATION NUMBER: US/10/170,205E
; CURRENT FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 40312
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 27511
; LENGTH: 260
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-170-205E-27511

Query Match      87.1%; Score 1332; DB 31; Length 260;
Best Local Similarity 96.9%; Pred. No. 2.2e-133;
Matches 249; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 MAESAETPSTGPEFESKYFEHGVRLPPFCRGMKEEIANFPVRPSDVWIVTPKSGTSL 60
DB 1 MAESAETPSTGPEFESKYFEHGVRLPPFCRGMKEEIANFPVRPSDVWIVTPKSGTSL 60
QY 61 LOEVVYLVSOQADPDEIGLWNIDQPLVLEYPOQGLDIIKELTSPRLIKSHLPYRFLPSD 120
DB 61 LOEVVYLVSOQADPDEIGLWNIDQPLVLEYPOQGLDIIKELTSPRLIKSHLPYRFLPSD 120
QY 121 LHNGDSKVIYMARNPKDLVVSYYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHGVQ 180
DB 121 LHNGDSKVIYMARNPKDLVVSYYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHGVQ 180
QY 181 EFWEHRMDSNVFLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLAETHECHQVLDQCCNA 240
DB 181 EFWEHRMDSNVFLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLAETHECHQVLDQCCNA 240
QY 241 EALPVGRVGLWKDIF 257
DB 241 EALPVGRAHCVFARKIF 257

RESULT 54
US-10-990-328-7684
; Sequence 7684, Application US/10990328
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; FILE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL001495
; CURRENT APPLICATION NUMBER: US/10/990,328
; CURRENT FILING DATE: 2004-11-17
; NUMBER OF SEQ ID NOS: 558824
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7684
; LENGTH: 260
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-990-328-7684

Query Match      87.1%; Score 1332; DB 39; Length 260;
Best Local Similarity 96.9%; Pred. No. 2.2e-133;
Matches 249; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 MAESAETPSTGPEFESKYFEHGVRLPPFCRGMKEEIANFPVRPSDVWIVTPKSGTSL 60
DB 1 MAESAETPSTGPEFESKYFEHGVRLPPFCRGMKEEIANFPVRPSDVWIVTPKSGTSL 60
QY 61 LOEVVYLVSOQADPDEIGLWNIDQPLVLEYPOQGLDIIKELTSPRLIKSHLPYRFLPSD 120
DB 61 LOEVVYLVSOQADPDEIGLWNIDQPLVLEYPOQGLDIIKELTSPRLIKSHLPYRFLPSD 120
QY 121 LHNGDSKVIYMARNPKDLVVSYYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHGVQ 180
DB 121 LHNGDSKVIYMARNPKDLVVSYYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHGVQ 180
QY 181 EFWEHRMDSNVFLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLAETHECHQVLDQCCNA 240
DB 181 EFWEHRMDSNVFLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLAETHECHQVLDQCCNA 240
QY 241 EALPVGRVGLWKDIF 257
DB 241 EALPVGRAHCVFARKIF 257
```

```
RESULT 53
US-10-932-348-406
; Sequence 406, Application US/10932348
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; FILE OF INVENTION: ENCODING HUMAN DRUG-METABOLIZING ENZYMES, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001481
; CURRENT APPLICATION NUMBER: US/10/932,348
; CURRENT FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 30431
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 406
; LENGTH: 260
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-932-348-406

Query Match      87.1%; Score 1332; DB 39; Length 260;
Best Local Similarity 96.9%; Pred. No. 2.2e-133;
Matches 249; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 MAESAETPSTGPEFESKYFEHGVRLPPFCRGMKEEIANFPVRPSDVWIVTPKSGTSL 60
DB 1 MAESAETPSTGPEFESKYFEHGVRLPPFCRGMKEEIANFPVRPSDVWIVTPKSGTSL 60
QY 61 LOEVVYLVSOQADPDEIGLWNIDQPLVLEYPOQGLDIIKELTSPRLIKSHLPYRFLPSD 120
DB 61 LOEVVYLVSOQADPDEIGLWNIDQPLVLEYPOQGLDIIKELTSPRLIKSHLPYRFLPSD 120
QY 121 LHNGDSKVIYMARNPKDLVVSYYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHGVQ 180
DB 121 LHNGDSKVIYMARNPKDLVVSYYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHGVQ 180
QY 181 EFWEHRMDSNVFLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLAETHECHQVLDQCCNA 240
DB 181 EFWEHRMDSNVFLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLAETHECHQVLDQCCNA 240
QY 241 EALPVGRVGLWKDIF 257
DB 241 EALPVGRAHCVFARKIF 257

RESULT 54
US-10-990-328-7684
; Sequence 7684, Application US/10990328
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; FILE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL001495
; CURRENT APPLICATION NUMBER: US/10/990,328
; CURRENT FILING DATE: 2004-11-17
; NUMBER OF SEQ ID NOS: 558824
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7684
; LENGTH: 260
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-990-328-7684

Query Match      87.1%; Score 1332; DB 39; Length 260;
Best Local Similarity 96.9%; Pred. No. 2.2e-133;
Matches 249; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 MAESAETPSTGPEFESKYFEHGVRLPPFCRGMKEEIANFPVRPSDVWIVTPKSGTSL 60
DB 1 MAESAETPSTGPEFESKYFEHGVRLPPFCRGMKEEIANFPVRPSDVWIVTPKSGTSL 60
QY 61 LOEVVYLVSOQADPDEIGLWNIDQPLVLEYPOQGLDIIKELTSPRLIKSHLPYRFLPSD 120
DB 61 LOEVVYLVSOQADPDEIGLWNIDQPLVLEYPOQGLDIIKELTSPRLIKSHLPYRFLPSD 120
QY 121 LHNGDSKVIYMARNPKDLVVSYYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHGVQ 180
DB 121 LHNGDSKVIYMARNPKDLVVSYYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHGVQ 180
QY 181 EFWEHRMDSNVFLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLAETHECHQVLDQCCNA 240
DB 181 EFWEHRMDSNVFLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLAETHECHQVLDQCCNA 240
QY 241 EALPVGRVGLWKDIF 257
DB 241 EALPVGRAHCVFARKIF 257
```



```
Db 61 LQEVVYLVSQGADPDEIGLNMNIDQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120
;
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-499-964-406

Query Match 87.1%; Score 1332; DB 39; Length 260;
Best Local Similarity 96.9%; Pred. No. 2.2e-133;
Matches 249; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 1 MAESEAETPTGPFESKYFEFHGVRLLPPFCRGKMEEIANFPVRPSDVWIVTPKSGTSL 60
Db 1 MAESEAETPTGPFESKYFEFHGVRLLPPFCRGKMEEIANFPVRPSDVWIVTPKSGTSL 60

Qy 61 LQEVVYLVSQGADPDEIGLNMNIDQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LQEVVYLVSQGADPDEIGLNMNIDQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120

Qy 121 LHNGDSKVIYARNPKDLVVSYYQFHRSLRTMSYRGTFQEFCCRPMNDKLGYSWFHVQ 180
Db 121 LHNGDSKVIYARNPKDLVVSYYQFHRSLRTMSYRGTFQEFCCRPMNDKLGYSWFHVQ 180

Qy 181 EFWEHRMDSNVFLFKYEDMHRDLVTMVQELARFLGVSCDKAQLALTEHCHQLVDQCCNA 240
Db 181 EFWEHRMDSNVFLFKYEDMHRDLVTMVQELARFLGVSCDKAQLALTEHCHQLVDQCCNA 240

Qy 241 EALPVGRGRVGLWKDIF 257
Db 241 EALPVGRAHCVFARKIF 257

RESULT 55
US-10-990-328A-7684
; Sequence 7684, Application US/10990328A
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL001495
; CURRENT APPLICATION NUMBER: US/10/990,328A
; CURRENT FILING DATE: 2004-11-17
; NUMBER OF SEQ ID NOS: 558824
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7684
; LENGTH: 260
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-990-328A-7684

Query Match 87.1%; Score 1332; DB 39; Length 260;
Best Local Similarity 96.9%; Pred. No. 2.2e-133;
Matches 249; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 1 MAESEAETPTGPFESKYFEFHGVRLLPPFCRGKMEEIANFPVRPSDVWIVTPKSGTSL 60
Db 1 MAESEAETPTGPFESKYFEFHGVRLLPPFCRGKMEEIANFPVRPSDVWIVTPKSGTSL 60

Qy 61 LQEVVYLVSQGADPDEIGLNMNIDQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LQEVVYLVSQGADPDEIGLNMNIDQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120

Qy 121 LHNGDSKVIYARNPKDLVVSYYQFHRSLRTMSYRGTFQEFCCRPMNDKLGYSWFHVQ 180
Db 121 LHNGDSKVIYARNPKDLVVSYYQFHRSLRTMSYRGTFQEFCCRPMNDKLGYSWFHVQ 180

Qy 181 EFWEHRMDSNVFLFKYEDMHRDLVTMVQELARFLGVSCDKAQLALTEHCHQLVDQCCNA 240
Db 181 EFWEHRMDSNVFLFKYEDMHRDLVTMVQELARFLGVSCDKAQLALTEHCHQLVDQCCNA 240

Qy 241 EALPVGRGRVGLWKDIF 257
Db 241 EALPVGRAHCVFARKIF 257

RESULT 56
US-60-499-964-406
; Sequence 406, Application US/60499964
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN DRUG-METABOLIZING ENZYMES, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001481
; CURRENT APPLICATION NUMBER: US/60/499,964
; CURRENT FILING DATE: 2003-09-04
; NUMBER OF SEQ ID NOS: 30431
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 406
```

```
; LENGTH: 260
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-499-964-406

Query Match 87.1%; Score 1332; DB 49; Length 260;
Best Local Similarity 96.9%; Pred. No. 2.2e-133;
Matches 249; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 1 MAESEAETPTGPFESKYFEFHGVRLLPPFCRGKMEEIANFPVRPSDVWIVTPKSGTSL 60
Db 1 MAESEAETPTGPFESKYFEFHGVRLLPPFCRGKMEEIANFPVRPSDVWIVTPKSGTSL 60

Qy 61 LQEVVYLVSQGADPDEIGLNMNIDQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LQEVVYLVSQGADPDEIGLNMNIDQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120

Qy 121 LHNGDSKVIYARNPKDLVVSYYQFHRSLRTMSYRGTFQEFCCRPMNDKLGYSWFHVQ 180
Db 121 LHNGDSKVIYARNPKDLVVSYYQFHRSLRTMSYRGTFQEFCCRPMNDKLGYSWFHVQ 180

Qy 181 EFWEHRMDSNVFLFKYEDMHRDLVTMVQELARFLGVSCDKAQLALTEHCHQLVDQCCNA 240
Db 181 EFWEHRMDSNVFLFKYEDMHRDLVTMVQELARFLGVSCDKAQLALTEHCHQLVDQCCNA 240

Qy 241 EALPVGRGRVGLWKDIF 257
Db 241 EALPVGRAHCVFARKIF 257

RESULT 57
US-10-932-348-407
; Sequence 407, Application US/10932348
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN DRUG-METABOLIZING ENZYMES, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001481
; CURRENT APPLICATION NUMBER: US/10/932,348
; CURRENT FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 30431
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 407
; LENGTH: 286
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-932-348-407

Query Match 82.8%; Score 1267; DB 39; Length 286;
Best Local Similarity 84.5%; Pred. No. 2.4e-126;
Matches 246; Conservative 7; Mismatches 26; Indels 12; Gaps 3;

Qy 1 MAESEAETPTGPFESKYFEFHGVRLLPPFCRGKMEEIANFPVRPSDVWIVTPKSGTSL 60
Db 1 MAESEAETPTGPFESKYFEFHGVRLLPPFCRGKMEEIANFPVRPSDVWIVTPKSGTSL 60

Qy 61 LQEVVYLVSQGADPDEIGLNMNIDQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LQEVVYLVSQGADPDEIGLNMNIDQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120

Qy 121 LHNGDSKVIYARNPKDLVVSYYQFHRSLRTMSYRGTFQEFCCRPMNDKLGYSWFHVQ 180
Db 121 LHNGDSKVIYARNPKDLVVSYYQFHRSLRTMSYRGTFQEFCCRPMNDKLGYSWFHVQ 180

Qy 181 EFWEHRMDSNVFLFKYEDMHRDLVTMVQELARFLGV-----SCDKAQLALTEHCHQL 233
Db 181 EFWEHRMDSNVFLFKYEDMHRDLVTMVQELARFLGV-----SCDKAQLALTEHCHQL 233

Qy 234 VDQCCNAEALPVGRGRVGLWKDIFTVSNNEKFDLVYKQMGKCDLTFDFYL 284
Db 237 -TASGPSSSLVCVSGRVGLWKDIFTVSNNEKFDLVYKQMGKCDLTFDFYL 286
```

```
RESULT 58
US-10-990-328-7685
; Sequence 7685, Application US/10990328
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
; FILE REFERENCE: CL001495
; CURRENT APPLICATION NUMBER: US/10/990,328
; NUMBER OF SEQ ID NOS: 558824
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7685
; LENGTH: 286
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-990-328-7685

Query Match      82.8%; Score 1267; DB 39; Length 286;
Best Local Similarity 84.5%; Pred. No. 2.4e-126;
Matches 246; Conservative 7; Mismatches 26; Indels 12; Gaps 3;

QY 1 MAESAETPSTPGEFESKYFEFGVRLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60
DB 1 MAESAETPSTPGEFESKYFEFGVRLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60
QY 61 LOEVVYLSQGADPDEIGLMNIDEOLPVLEYQPQGLDIIKELTSRLIKSHLPYRFLPSD 120
DB 61 LOEVVYLSQGADPDEIGLMNIDEOLPVLEYQPQGLDIIKELTSRLIKSHLPYRFLPSD 120
QY 121 LHNGDSKVIYMARNP KD LVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180
DB 121 LHNGDSKVIYMARNP KD LVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180
QY 181 EFWEHRMDSNVLFKYEDMHRD-LVTWVEQLARFLGV-----SCDKAQLEALTEHCQOL 233
DB 181 EFWEHRMDSNVLFKYEDMHRD-LVTWVEQLARFLGV-----SCDKAQLEALTEHCQOL 233
QY 234 VDQCCNAEALPVGRGVLWKDIFTVSMNEKFDLVYKQMGKCDLTDFDYL 284
DB 234 VDQCCNAEALPVGRGVLWKDIFTVSMNEKFDLVYKQMGKCDLTDFDYL 284
QY 237 -TASGPSSSLVCVSRGVLWKDIFTVSMNEKFDLVYKQMGKCDLTDFDYL 286
DB 237 -TASGPSSSLVCVSRGVLWKDIFTVSMNEKFDLVYKQMGKCDLTDFDYL 286

RESULT 59
US-10-990-328A-7685
; Sequence 7685, Application US/10990328A
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
; FILE REFERENCE: CL001495
; CURRENT APPLICATION NUMBER: US/10/990,328A
; CURRENT FILING DATE: 2004-11-17
; NUMBER OF SEQ ID NOS: 558824
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7685
; LENGTH: 286
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-990-328A-7685

Query Match      82.8%; Score 1267; DB 39; Length 286;
Best Local Similarity 84.5%; Pred. No. 2.4e-126;
Matches 246; Conservative 7; Mismatches 26; Indels 12; Gaps 3;

QY 1 MAESAETPSTPGEFESKYFEFGVRLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60
DB 1 MAESAETPSTPGEFESKYFEFGVRLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60
QY 61 LOEVVYLSQGADPDEIGLMNIDEOLPVLEYQPQGLDIIKELTSRLIKSHLPYRFLPSD 120
DB 61 LOEVVYLSQGADPDEIGLMNIDEOLPVLEYQPQGLDIIKELTSRLIKSHLPYRFLPSD 120
QY 121 LHNGDSKVIYMARNP KD LVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180
DB 121 LHNGDSKVIYMARNP KD LVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180
QY 181 EFWEHRMDSNVLFKYEDMHRD-LVTWVEQLARFLGV-----SCDKAQLEALTEHCQOL 233
DB 181 EFWEHRMDSNVLFKYEDMHRD-LVTWVEQLARFLGV-----SCDKAQLEALTEHCQOL 233
QY 234 VDQCCNAEALPVGRGVLWKDIFTVSMNEKFDLVYKQMGKCDLTDFDYL 284
DB 234 VDQCCNAEALPVGRGVLWKDIFTVSMNEKFDLVYKQMGKCDLTDFDYL 284
QY 237 -TASGPSSSLVCVSRGVLWKDIFTVSMNEKFDLVYKQMGKCDLTDFDYL 286
DB 237 -TASGPSSSLVCVSRGVLWKDIFTVSMNEKFDLVYKQMGKCDLTDFDYL 286

RESULT 60
US-10-990-964-407
; Sequence 407, Application US/60499964
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN DRUG-METABOLIZING ENZYMES, METHODS OF
; FILE REFERENCE: CL001481
; CURRENT APPLICATION NUMBER: US/60/499,964
; CURRENT FILING DATE: 2003-09-04
; NUMBER OF SEQ ID NOS: 30431
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 407
; LENGTH: 286
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-990-964-407

Query Match      82.8%; Score 1267; DB 49; Length 286;
Best Local Similarity 84.5%; Pred. No. 2.4e-126;
Matches 246; Conservative 7; Mismatches 26; Indels 12; Gaps 3;

QY 1 MAESAETPSTPGEFESKYFEFGVRLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60
DB 1 MAESAETPSTPGEFESKYFEFGVRLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60
QY 61 LOEVVYLSQGADPDEIGLMNIDEOLPVLEYQPQGLDIIKELTSRLIKSHLPYRFLPSD 120
DB 61 LOEVVYLSQGADPDEIGLMNIDEOLPVLEYQPQGLDIIKELTSRLIKSHLPYRFLPSD 120
QY 121 LHNGDSKVIYMARNP KD LVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180
DB 121 LHNGDSKVIYMARNP KD LVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180
QY 181 EFWEHRMDSNVLFKYEDMHRD-LVTWVEQLARFLGV-----SCDKAQLEALTEHCQOL 233
DB 181 EFWEHRMDSNVLFKYEDMHRD-LVTWVEQLARFLGV-----SCDKAQLEALTEHCQOL 233
QY 234 VDQCCNAEALPVGRGVLWKDIFTVSMNEKFDLVYKQMGKCDLTDFDYL 284
DB 234 VDQCCNAEALPVGRGVLWKDIFTVSMNEKFDLVYKQMGKCDLTDFDYL 284
QY 237 -TASGPSSSLVCVSRGVLWKDIFTVSMNEKFDLVYKQMGKCDLTDFDYL 286
DB 237 -TASGPSSSLVCVSRGVLWKDIFTVSMNEKFDLVYKQMGKCDLTDFDYL 286

RESULT 61
PCT-US02-11654-172
; Sequence 172, Application PC/TUS0211654
; GENERAL INFORMATION:
; APPLICANT: Agenays, Inc.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
```

/	TITLE OF INVENTION:	Nucleic Acids and Corresponding Proteins Useful			
/	TITLE OF INVENTION:	in the Detection and Treatment of Various Cancers			
/	FILE REFERENCE:	51158-20040.40			
/	CURRENT APPLICATION NUMBER:	PCT/US02/11654			
/	CURRENT FILING DATE:	2002-04-10			
/	PRIOR APPLICATION NUMBER:	US 60/283,112			
/	PRIOR FILING DATE:	2001-04-10			
/	PRIOR APPLICATION NUMBER:	US 60/282,739			
/	PRIOR FILING DATE:	2001-04-10			
/	PRIOR APPLICATION NUMBER:	US 60/286,630			
/	PRIOR FILING DATE:	2001-04-25			
/	NUMBER OF SEQ ID NOS:	708			
/	SOFTWARE:	FastSeq for Windows Version 4.0			
/	SEQ ID NO 172				
/	LENGTH:	218			
/	TYPE:	PRT			
/	ORGANISM:	Homo sapiens			
/	PCT-US02-11654-172				
	Query Match	65.5%; Score 1002; DB 1; Length 218;			
	Best Local Similarity	96.4%; Pred. No. 4.5e-98;			
	Matches	186; Conservative 0; Mismatches 1; Indels 6; Gaps 1;			
Qy	92	PQGLDIIIEKLSPLIKSHLPYRFLPSPDLHNGDSKVITYMARNPKDIWVSYQPFRSLRT	151		
Dd	32	PSP-----ELTSPRLIKSHLPYRFLPSPDLHNGDSKVITYMARNPKDIWVSYQPFRSLRT	85		
Qy	152	MSVRGTGFQBFRCRRFMNDKLGYSWPBHVQEFWEHRMDSNVLFKYEDMHRDLVTMVQOLA	211		
Dd	86	MSVRGTGFQBFRCRRFMNDKLGYSWFEBHVQEFWEHRMDSNVLFKYEDMHRDLVTMVQOLA	145		
Qy	212	RFLGVSCDKAQAEALTEHCHQLVDQCNAEALPVGRGRVGLWKDIFTVSMNEKFDLVYKQ	271		
Dd	146	RFLGVSCDKAQAEALTEHCHQLVDQCNAEALPVGRGRVGLWKDIFTVSMNEKFDLVYKQ	205		
Qy	272	KMGKCDLTDFYL	284		
Dd	206	KMGKCDLTDFYL	218		
RESULT 62					
US-10-121-024B-172					
/	Sequence 172,	Application US/10121024B			
/	GENERAL INFORMATION:				
/	APPLICANT:	Agensys, Inc.			
/	APPLICANT:	Jakobovits, Aya			
/	APPLICANT:	Challita-Eid, Fia M.			
/	APPLICANT:	Faris, Mary			
/	APPLICANT:	Ge, Wangmao			
/	APPLICANT:	Hubert, Rene S.			
/	APPLICANT:	Morrison, Karen Jane Meyrick			
/	APPLICANT:	Morrison, Robert Kendall			
/	APPLICANT:	Raitano, Arthur B.			
/	TITLE OF INVENTION:	Nucleic Acids and Corresponding Proteins			
/	TITLE OF INVENTION:	Useful in the Detection and Treatment of Various Cancers			
/	FILE REFERENCE:	51158-20040.00			
/	CURRENT APPLICATION NUMBER:	US/10/121,024B			
/	CURRENT FILING DATE:	2002-04-10			
/	PRIOR APPLICATION NUMBER:	US60/283,112			
/	PRIOR FILING DATE:	2001-04-10			
/	PRIOR APPLICATION NUMBER:	US60/282,739			
/	PRIOR FILING DATE:	2001-04-10			
/	PRIOR APPLICATION NUMBER:	US60/286,630			
/	PRIOR FILING DATE:	2001-04-25			
/	NUMBER OF SEQ ID NOS:	708			
/	SOFTWARE:	FastSeq for Windows Version 4.0			
/	SEQ ID NO 172				
/	LENGTH:	218			
/	TYPE:	PRT			
/	ORGANISM:	Homo sapiens			
/	US-10-121-024B-172				
	Query Match	65.5%; Score 1002; DB 31; Length 218;			

```
US-11-154-298-172
; Sequence 172, Application US/11154298
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: METHODS OF INDUCING AN IMMUNE RESPONSE
; FILE REFERENCE: 51158-20040.01
; CURRENT APPLICATION NUMBER: US/11/154,298
; CURRENT FILING DATE: 2005-06-15
; PRIOR FILING DATE: 2002-04-10
; PRIOR FILING DATE: 2002-04-10
; PRIOR FILING DATE: 2002-04-10
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 172
; LENGTH: 218
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-154-298-172

Query Match      65.5%; Score 1002; DB 41; Length 218;
Best Local Similarity 96.4%; Pred. No. 4.5e-98;
Matches 186; Conservative 0; Mismatches 1; Indels 6; Gaps 1;

QY 92 POPGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARPNKDLVVSYYQFHRSLRT 151
Db 32 PSP-----ELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARPNKDLVVSYYQFHRSLRT 85

QY 152 MSYRGTFOEFCRRFMNDKLGYSWFHVEHQRMDNSVLFKLYEDMHRDLVTWVEOLA 211
Db 86 MSYRGTFOEFCRRFMNDKLGYSWFHVEHQRMDNSVLFKLYEDMHRDLVTWVEOLA 145

QY 212 RFLGVSCDKAQLEALTEHCHQLVDQCCNAEALPVGRGRVGLWKDIFTVSMNEKFDLVYKQ 271
Db 146 RFLGVSCDKAQLEALTEHCHQLVDQCCNAEALPVGRGRVGLWKDIFTVSMNEKFDLVYKQ 205

QY 272 KMGKCDLTFDFYL 284
Db 206 KMGKCDLTFDFYL 218

US-11-154-298-172

Query Match      65.5%; Score 1002; DB 41; Length 218;
Best Local Similarity 96.4%; Pred. No. 4.5e-98;
Matches 186; Conservative 0; Mismatches 1; Indels 6; Gaps 1;

QY 92 POPGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARPNKDLVVSYYQFHRSLRT 151
Db 32 PSP-----ELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARPNKDLVVSYYQFHRSLRT 85

QY 152 MSYRGTFOEFCRRFMNDKLGYSWFHVEHQRMDNSVLFKLYEDMHRDLVTWVEOLA 211
Db 86 MSYRGTFOEFCRRFMNDKLGYSWFHVEHQRMDNSVLFKLYEDMHRDLVTWVEOLA 145

QY 212 RFLGVSCDKAQLEALTEHCHQLVDQCCNAEALPVGRGRVGLWKDIFTVSMNEKFDLVYKQ 271
Db 146 RFLGVSCDKAQLEALTEHCHQLVDQCCNAEALPVGRGRVGLWKDIFTVSMNEKFDLVYKQ 205

QY 272 KMGKCDLTFDFYL 284
Db 206 KMGKCDLTFDFYL 218

US-11-155-906-172
; Sequence 172, Application US/11155906
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: METHODS OF INDUCING AN IMMUNE RESPONSE
; FILE REFERENCE: 51158-20040.02
; CURRENT APPLICATION NUMBER: US/11/155,906
; CURRENT FILING DATE: 2005-06-17
; PRIOR FILING DATE: 2002-04-10
; PRIOR FILING DATE: 2002-04-10
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 172
; LENGTH: 218
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-155-906-172

Query Match      65.5%; Score 1002; DB 41; Length 218;
Best Local Similarity 96.4%; Pred. No. 4.5e-98;
Matches 186; Conservative 0; Mismatches 1; Indels 6; Gaps 1;

QY 92 POPGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARPNKDLVVSYYQFHRSLRT 151
Db 32 PSP-----ELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARPNKDLVVSYYQFHRSLRT 85

QY 152 MSYRGTFOEFCRRFMNDKLGYSWFHVEHQRMDNSVLFKLYEDMHRDLVTWVEOLA 211
Db 86 MSYRGTFOEFCRRFMNDKLGYSWFHVEHQRMDNSVLFKLYEDMHRDLVTWVEOLA 145

QY 212 RFLGVSCDKAQLEALTEHCHQLVDQCCNAEALPVGRGRVGLWKDIFTVSMNEKFDLVYKQ 271
Db 146 RFLGVSCDKAQLEALTEHCHQLVDQCCNAEALPVGRGRVGLWKDIFTVSMNEKFDLVYKQ 205

QY 272 KMGKCDLTFDFYL 284
Db 206 KMGKCDLTFDFYL 218

US-11-156-231-172
; Sequence 172, Application US/11156231
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: METHODS OF INDUCING AN IMMUNE RESPONSE
; FILE REFERENCE: 51158-20040.03
; CURRENT APPLICATION NUMBER: US/11/156,231
; CURRENT FILING DATE: 2005-06-17
; PRIOR APPLICATION NUMBER: US10/121,024
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 172
; LENGTH: 218
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-156-231-172

Query Match      65.5%; Score 1002; DB 41; Length 218;
Best Local Similarity 96.4%; Pred. No. 4.5e-98;
Matches 186; Conservative 0; Mismatches 1; Indels 6; Gaps 1;

QY 92 POPGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARPNKDLVVSYYQFHRSLRT 151
Db 32 PSP-----ELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARPNKDLVVSYYQFHRSLRT 85

QY 152 MSYRGTFOEFCRRFMNDKLGYSWFHVEHQRMDNSVLFKLYEDMHRDLVTWVEOLA 211
Db 86 MSYRGTFOEFCRRFMNDKLGYSWFHVEHQRMDNSVLFKLYEDMHRDLVTWVEOLA 145

QY 212 RFLGVSCDKAQLEALTEHCHQLVDQCCNAEALPVGRGRVGLWKDIFTVSMNEKFDLVYKQ 271
Db 146 RFLGVSCDKAQLEALTEHCHQLVDQCCNAEALPVGRGRVGLWKDIFTVSMNEKFDLVYKQ 205

QY 272 KMGKCDLTFDFYL 284
Db 206 KMGKCDLTFDFYL 218
```

Db 86 MSYRGTFEFCRRFNDKLGSGSFEHVQEFWEHRMDSNVLFKVEDMHRDLVTWVEQLA 145
Qy 212 RFLGVSCDKAQLALETCHQLVDQCCNAEALPVGRGRVGLWKDIFTVSMNEKFDLVYKQ 271
Db 146 RFLGVSCDKAQLALETCHQLVDQCCNAEALPVGRGRVGLWKDIFTVSMNEKFDLVYKQ 205
Qy 272 KMGKCDLTFDFYL 284
Db 206 KMGKCDLTFDFYL 218

RESULT 67

PCT-US02-11654-173
; Sequence 173, Application PC/TUS0211654
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins Useful
; FILE REFERENCE: 51158-20040.40
; CURRENT APPLICATION NUMBER: PCT/US02/11654
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US 60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 173
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-11654-173

Query Match 56.7%; Score 867.5; DB 1; Length 171;
Best Local Similarity 59.9%; Pred. No. 8.6e-84;
Matches 170; Conservative 1; Mismatches 0; Indels 113; Gaps 1;
Qy 1 MAESEAETPTGGEFESKYFEHGVRLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60
Db 1 MAESEAETPTGGEFESKYFEHGVRLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 56
Qy 61 LOEVVYLVSGADPDEIGLWNIDEQLPVLEYPPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 57 -----
Qy 121 LHNGDSKVIYARNPKDLVVSYYQFHRSLRTMSYRGTFQFCRRFNDKLGSGSFEHVQ 180
Db 57 -----VGYSWFHVQ 67
Qy 181 EFWHRMDSNVLFKVEDMHRDLVTWVEQLARFLGVSCDKAQLALETCHQLVDQCCNA 240
Db 68 EFWHRMDSNVLFKVEDMHRDLVTWVEQLARFLGVSCDKAQLALETCHQLVDQCCNA 127
Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284
Db 128 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 171

RESULT 68

US-10-121-024B-173
; Sequence 173, Application US/10121024B
; GENERAL INFORMATION:

; APPLICANT: Agensys, Inc.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins Useful
; FILE REFERENCE: 51158-20040.00
; CURRENT APPLICATION NUMBER: US/10/121,024B
; CURRENT FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 173
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-121-024B-173

Query Match 56.7%; Score 867.5; DB 31; Length 171;
Best Local Similarity 59.9%; Pred. No. 8.6e-84;
Matches 170; Conservative 1; Mismatches 0; Indels 113; Gaps 1;
Qy 1 MAESEAETPTGGEFESKYFEHGVRLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60
Db 1 MAESEAETPTGGEFESKYFEHGVRLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 56
Qy 61 LOEVVYLVSGADPDEIGLWNIDEQLPVLEYPPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 57 -----
Qy 121 LHNGDSKVIYARNPKDLVVSYYQFHRSLRTMSYRGTFQFCRRFNDKLGSGSFEHVQ 180
Db 57 -----VGYSWFHVQ 67
Qy 181 EFWHRMDSNVLFKVEDMHRDLVTWVEQLARFLGVSCDKAQLALETCHQLVDQCCNA 240
Db 68 EFWHRMDSNVLFKVEDMHRDLVTWVEQLARFLGVSCDKAQLALETCHQLVDQCCNA 127
Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284
Db 128 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 171

RESULT 69

US-10-121-024C-173
; Sequence 173, Application US/10121024C
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins Useful
; FILE REFERENCE: 51158-20040.00
; CURRENT APPLICATION NUMBER: US/10/121,024C
; CURRENT FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739


```
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Raitano, Robert B.
; TITLE OF INVENTION: METHODS OF INDUCING AN IMMUNE RESPONSE
; FILE REFERENCE: 51158-20040.02
; CURRENT APPLICATION NUMBER: US/11/155,906
; CURRENT FILING DATE: 2005-06-17
; PRIOR APPLICATION NUMBER: US10/121,024
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 173
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-11-155-906-173

Query Match      56.7%; Score 867.5; DB 41; Length 171;
Best Local Similarity 59.9%; Pred. No. 8.6e-84;
Matches 170; Conservative 1; Mismatches 0; Indels 113; Gaps 1;

QY 1 MAESAETPTSGEFESKYFEFGHVRLLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60
Db 1 MAESAETPTSGEFESKYFEFGHVRLLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60
QY 61 LOEVVYLVSQGADPEIGLWNIDEQLPVLEYQPQGLDIKELTSPRLIKSHLPYRFLPSD 120
Db 57 -----
QY 121 LHNGDSKVIYARNPKDLVSVYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFHVVQ 180
Db 57 -----
QY 181 EFWHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLAETHECHQLVDQCCNA 240
Db 68 EFWHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLAETHECHQLVDQCCNA 127
QY 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284
Db 128 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 171

RESULT 73
US-11-156-231-173
; Sequence 173, Application US/11156231
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: METHODS OF INDUCING AN IMMUNE RESPONSE
; FILE REFERENCE: 51158-20040.03
; CURRENT APPLICATION NUMBER: US/11/156,231
; CURRENT FILING DATE: 2005-06-17
; PRIOR APPLICATION NUMBER: US10/121,024
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
```

```
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 173
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-11-156-231-173

Query Match      56.7%; Score 867.5; DB 41; Length 171;
Best Local Similarity 59.9%; Pred. No. 8.6e-84;
Matches 170; Conservative 1; Mismatches 0; Indels 113; Gaps 1;

QY 1 MAESAETPTSGEFESKYFEFGHVRLLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60
Db 1 MAESAETPTSGEFESKYFEFGHVRLLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60
QY 61 LOEVVYLVSQGADPEIGLWNIDEQLPVLEYQPQGLDIKELTSPRLIKSHLPYRFLPSD 120
Db 57 -----
QY 121 LHNGDSKVIYARNPKDLVSVYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFHVVQ 180
Db 57 -----
QY 181 EFWHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLAETHECHQLVDQCCNA 240
Db 68 EFWHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLAETHECHQLVDQCCNA 127
QY 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284
Db 128 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 171

RESULT 74
PCT-US01-41986-2
; Sequence 2, Application PC/TUS0141986
; GENERAL INFORMATION:
; APPLICANT: Trustees of Boston University
; TITLE OF INVENTION: A NEURAL SPECIFIC CYTOSOLIC
; TITLE OF INVENTION: SULFOTRANSFERASE FOR DRUG SCREENING
; FILE REFERENCE: 13594-007W01
; CURRENT APPLICATION NUMBER: PCT/US01/41986
; CURRENT FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: 60/229,929
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Rattus sp.
; PCT-US01-41986-2

Query Match      56.2%; Score 860; DB 1; Length 160;
Best Local Similarity 98.1%; Pred. No. 5e-83;
Matches 157; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 125 DSKVIYARNPKDLVSVYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFHVVQEFWE 184
Db 1 DSKVIYARNPKDLVSVYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFHVVQEFWE 60
QY 185 HRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLAETHECHQLVDQCCNAEALP 244
Db 61 HRMDANVLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLAETHECHQLVDQCCNAEALP 120
QY 245 VGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284
Db 121 VGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 160
```


RESULT 75
US-10-170-205E-27771
; Sequence 27771, Application US/10170205E
; GENERAL INFORMATION:
; APPLICANT: ADAMS, Mark
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN
; TITLE OF INVENTION: CAPTURE AGENTS, AND USES THEREOF
; FILE REFERENCE: CL001381
; CURRENT APPLICATION NUMBER: US/10/170,205E
; CURRENT FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 40312
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 27771
; LENGTH: 194
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-170-205E-27771

Query Match 52.5%; Score 804; DB 31; Length 194;
Best Local Similarity 91.0%; Pred. No. 6.7e-77;
Matches 151; Conservative 1; Mismatches 8; Indels 6; Gaps 1;

Qy 92 POPGLDIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARPKDLVVSYYQFHRSLRT 151
Db 32 PSP-----ELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARPKDLVVSYYQFHRSLRT 85

Qy 152 MSYRGTFQFCRRFWMNDKLGYSWFQEHVQEFWEHRMDSNVLFKLYEDMHRDLVTMVEOLA 211
Db 86 MSYRGTFQFCRRFWMNDKLGYSWFQEHVQEFWEHRMDSNVLFKLYEDMHRDLVTMVEOLA 145

Qy 212 RFLGVSCDKAQLEALTEHCHQLVDQCNAEALPVGRGRVGLWKDIF 257
Db 146 RFLGVSCDKAQLEALTEHCHQLVDQCNAEALPVGRAHCVFARKIF 191

RESULT 76
US-09-791-537-24613
; Sequence 24613, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 24613
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-791-537-24613

Query Match 32.4%; Score 495; DB 27; Length 298;
Best Local Similarity 35.5%; Pred. No. 1.8e-43;
Matches 104; Conservative 54; Mismatches 117; Indels 18; Gaps 3;

Qy 6 AETPSTPGEFESKYFEFHGRVLRPPFCRGKMBEIANFPVRPSDVWIVTPYKSGTSLQEVV 65
Db 2 AQNPSNMEPLRKLPLVPVKGIPLIKVFATMEQLQNFATWPDVLISTYKSGTNNWSEIM 61

Qy 66 YLVSGADPDEIGLMNIDEQLPVLEYP-----QPLGDIKELTSPRLIKSHLPYRFLPSDL 121
Db 62 DMIVQGGKLDKGRAPVYARIPFLEFSCGVPVPGLETKETPAPRIIKTHLPLSLPOSTL 121

Qy 122 HNGDSKVIYMARPKDLVVSYYQFHRSLRTMSYRGTFQFCRRFWMNDKLGYSWFQEHVQ 181
Db 122 LDQIKVIYARNKDVVSYNFYKMAKLHPDPGTWESFLENFMDGKVSYSWYQHVKE 181

Qy 182 FWEHRMDSNVLFKLYEDMHRDLVTMVEQLARFLGVSCDKAQLEALTEHC---HQLVDQCC 238
Db 182 WWELRRTHPVLVLYFYEDMKENPKREIKKILFELGRSLPEETVDLIVVHTSPKKMKENPMA 241

Qy 239 NAEALP-----VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 242 NVTTIPTVMDHTIYPPFMRKGTIGDWKNTFTVAQSEHFDHAHYAKLMTGCDFTF 294

RESULT 78
US-09-791-537-66629
; Sequence 66629, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 66629
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-791-537-66629

Query Match 32.4%; Score 495; DB 27; Length 298;
Best Local Similarity 35.5%; Pred. No. 1.8e-43;
Matches 104; Conservative 54; Mismatches 117; Indels 18; Gaps 3;

Qy 6 AETPSTPGEFESKYFEFHGRVLRPPFCRGKMBEIANFPVRPSDVWIVTPYKSGTSLQEVV 65
Db 2 AQNPSNMEPLRKLPLVPVKGIPLIKVFATMEQLQNFATWPDVLISTYKSGTNNWSEIM 61

Qy 66 YLVSGADPDEIGLMNIDEQLPVLEYP-----QPLGDIKELTSPRLIKSHLPYRFLPSDL 121
Db 62 DMIVQGGKLDKGRAPVYARIPFLEFSCGVPVPGLETKETPAPRIIKTHLPLSLPOSTL 121

Qy 122 HNGDSKVIYMARPKDLVVSYYQFHRSLRTMSYRGTFQFCRRFWMNDKLGYSWFQEHVQ 181
Db 122 LDQIKVIYARNKDVVSYNFYKMAKLHPDPGTWESFLENFMDGKVSYSWYQHVKE 181

Qy 182 FWEHRMDSNVLFKLYEDMHRDLVTMVEQLARFLGVSCDKAQLEALTEHC---HQLVDQCC 238

Db 182 WWELRRTHPVLVLYFYEDMKENPKREIKKILFELGRSLPEETVDLIVVHTSPKKMKENPMA 241

Qy 239 NAEALP-----VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280

Db 242 NVTTIPTVMDHTIYPPFMRKGTIGDWKNTFTVAQSEHFDHAHYAKLMTGCDFTF 294

RESULT 77
US-09-791-537-26347
; Sequence 26347, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 26347
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-791-537-26347

Query Match 32.4%; Score 495; DB 27; Length 298;
Best Local Similarity 35.5%; Pred. No. 1.8e-43;
Matches 104; Conservative 54; Mismatches 117; Indels 18; Gaps 3;

Qy 6 AETPSTPGEFESKYFEFHGRVLRPPFCRGKMBEIANFPVRPSDVWIVTPYKSGTSLQEVV 65
Db 2 AQNPSNMEPLRKLPLVPVKGIPLIKVFATMEQLQNFATWPDVLISTYKSGTNNWSEIM 61

Qy 66 YLVSGADPDEIGLMNIDEQLPVLEYP-----QPLGDIKELTSPRLIKSHLPYRFLPSDL 121
Db 62 DMIVQGGKLDKGRAPVYARIPFLEFSCGVPVPGLETKETPAPRIIKTHLPLSLPOSTL 121

Qy 122 HNGDSKVIYMARPKDLVVSYYQFHRSLRTMSYRGTFQFCRRFWMNDKLGYSWFQEHVQ 181
Db 122 LDQIKVIYARNKDVVSYNFYKMAKLHPDPGTWESFLENFMDGKVSYSWYQHVKE 181

Qy 182 FWEHRMDSNVLFKLYEDMHRDLVTMVEQLARFLGVSCDKAQLEALTEHC---HQLVDQCC 238
Db 182 WWELRRTHPVLVLYFYEDMKENPKREIKKILFELGRSLPEETVDLIVVHTSPKKMKENPMA 241

Qy 239 NAEALP-----VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 242 NVTTIPTVMDHTIYPPFMRKGTIGDWKNTFTVAQSEHFDHAHYAKLMTGCDFTF 294

RESULT 78
US-09-791-537-66629
; Sequence 66629, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 66629
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-791-537-66629

Query Match 32.0%; Score 489; DB 27; Length 291;
Best Local Similarity 36.7%; Pred. No. 7.7e-43;
Matches 101; Conservative 53; Mismatches 103; Indels 18; Gaps 3;
QY 24 GVRPPFCRGMEEIANFVRPDSVMIIVTPYKSGTSLQLQEVVYLVSGQADPDEIGLMNID 83
DB 13 GIPLIKYPAETMBQQLNQFTAMPDDVLISTYKSGTNWSEIMDMYQGGKDKCGRAPVY 72
QY 84 EQLPVLEYP---QGLDIIKELTSRLIKSHLPYRFLPSDLHGDSKVIYMARPKDLV 139
DB 73 ARIPLEPSCGVPGLTKETAPRIKTHLPISLLPQSLDQIKIYIYARNAKDV 132
QY 140 VSYQFHRSLRMSYRGTFQFCRRFMDKLGYSWFEHVDHMDSNVLFKLYEDM 199
DB 133 VSYNYFYMAKLHPDPGGTWESFLENFMGDKVSYGSYQVORXEWELRTHPVLIFYEDM 192
QY 200 HRDLVTMVEQLARFLVSGCDKAQLEALTEHC---HQLVDQCCNAEALP-----V 245
DB 193 KENPKREIKLLEFLGRSLPEETVDLIVHTTSFKKMKENPMANYTTIPTVMDHTIYFPM 252
QY 246 GRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
DB 253 RKGTTGDWKNFTVAQSEHFAHAKLMTGCDFTF 287

RESULT 79

US-10-072-012-667
Sequence 667, Application US/10072012
GENERAL INFORMATION:
APPLICANT: Tchernev, Velizar
APPLICANT: Spytek, Kimberly
APPLICANT: Zernhusen, Bryan
APPLICANT: Patturajan, Meera
APPLICANT: Shinkets, Richard
APPLICANT: Li, Li
APPLICANT: Gangolli, Esha
APPLICANT: Padigaru, Muralidhara
APPLICANT: Anderson, David W.
APPLICANT: Rastelli, Luca
APPLICANT: Miller, Charles E.
APPLICANT: Gerlach, Valerie
APPLICANT: Taupier Jr, Raymond J.
APPLICANT: Gusev, Vladimir Y.
APPLICANT: Coleman, Steven D.
APPLICANT: Wolenc, Adam R.
APPLICANT: Pena, Carol E. A.
APPLICANT: Furtak, Katarzyna
APPLICANT: Grosse, William M.
APPLICANT: Alsobrook II, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-258
CURRENT APPLICATION NUMBER: US/10/072, 012
CURRENT FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: 60/265,102
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/265,514
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,517
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,412
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,395
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/266,406
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/266,767
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: 60/267,057
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/266,975

PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/267,459
PRIOR FILING DATE: 2001-02-08
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1391
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 667
LENGTH: 296
TYPE: PRT
ORGANISM: Mus musculus
US-10-072-012-667

Query Match 31.4%; Score 480; DB 30; Length 296;

Best Local Similarity 35.6%; Pred. No. 7.3e-42;
Matches 103; Conservative 57; Mismatches 107; Indels 22; Gaps 5;
QY 16 ESKYFEFHVRLPFCRCGMEEIANFVRPDSVMIIVTPYKSGTSLQLQEVVYLVSGADPD 75
DB 10 QTKLKEVAGIPLQAPTVDNWRQIQTFEAKPDDLLICTYKSGTTWQIEIVDMIEQNGDVE 69
QY 76 EIGLMNIDEQLPVLEY---POP-GLDIIKELTSRLIKSHLPYRFLPSDLHGDSKVIY 131
DB 70 KCRRTIIQHRHFFIEWARPPQPSGVDKANEMAPRILATHLPQTQLLPPSFWTNCKFLYV 129
QY 132 ARNPKDLVSYVYQFHRSLRMSYRGTFQFCRRFMDKLGYSWFEHVDHMDSNV 191
DB 130 ARNAKDCWVSYHYHFRMSQVLPEGTWDEYFETTINGKVSWSFHDHVGHWERDKYQI 189
QY 192 LFLKYEDMHRDLVTMVEQLARFLGVS-----CDKAQLEALTEHCQLVDOCCNAEALP-- 244
DB 190 LFLFYEDMKRNPKHEIQKVMQFMGNLDEDVVDKIVLETSFEKMKKE--NPMNTRSTAPKS 247
QY 245 -----VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL 284
DB 248 ILDOSISPFMRKGTGVDWKNHFTVAQNERFDEIYKQKMGRTSLNFSMEL 296

RESULT 80

US-60-212-357-126
Sequence 126, Application US/60212357
GENERAL INFORMATION:
APPLICANT: Beasley, Ellen
TITLE OF INVENTION: ISOLATED HUMAN PHASE II
TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, NUCLEIC ACID MOLECULES ENCODING
TITLE OF INVENTION: HUMAN DRUG-METABOLIZING PHASE II PROTEINS, AND USES THEREOF
FILE REFERENCE: CL000675
CURRENT APPLICATION NUMBER: US/60/212,357
CURRENT FILING DATE: 2000-06-19
NUMBER OF SEQ ID NOS: 237
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 126
LENGTH: 592
TYPE: PRT
ORGANISM: HUMAN
US-60-212-357-126

Query Match 31.4%; Score 480; DB 47; Length 592;

Best Local Similarity 33.7%; Pred. No. 2e-41;
Matches 112; Conservative 54; Mismatches 112; Indels 54; Gaps 8;
QY 3 ESEATPSTP-----GEFESKYFEHGVRL-----PP--FCRG-----K 34
DB 265 EGRQPTPTPTTILSPCRQGENRSGRELNNMELIODTSRPPLEYKGVPLIKYFAEA 324
QY 35 MEETANFVRPDSVMIIVTPYKSGTSLQLQEVVYLVSGQADPDEIGLMNIDEQLPVLEYPOP 94
DB 325 LGPLQSFQAPDDLLINTYKSGTTWVSQILDMIVQGDLEKCNRAPIYVRVPFLEVNDP 384
QY 95 ----GLDIIKELTSRLIKSHLPYRFLPSDLHGDSKVIYMARPKDLVSYVYQFHRSLR 150
DB 385 GEPSGLETIKDTPPRLIKSHLPALPLPQTLDDQKVVYVARNPKDVAVSYHHFRMEK 444
QY 151 TMSYRGTFQFCRRFMDKLGYSWFEHVDHMDSNVLFKLYEDMHRDLVTMVEQL 210

Db 445 AHPEPGTWDSELEKFMAGEVSYGSHYQHVQEWELSRTHPVLYLFYEDMKENPKREIQKI 504
Qy 211 ARFLGVSCDKAQLEALTEHC-----HQLVDOCCNABALP-VGRGRVGL 252
Db 505 LEFVGRSLPEETMDPMVQHTSFKEKKNPMNTYTTVPQELMDHSIS-----PPMRKGMAGD 560
Qy 253 WKDIFTVMNEKFDLVYKQKMGKCDLTFDFYL 284
Db 561 WKTFTVAQNERFDADYAEKMGAGCSLSRFE 592

RESULT 81
US-60-229-512-201
; Sequence 201, Application US/60229512
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PHASE II
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, NUCLEIC ACID MOLECULES ENCODING
; TITLE OF INVENTION: HUMAN DRUG-METABOLIZING PHASE II PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000770
; CURRENT APPLICATION NUMBER: US/60/229,512
; CURRENT FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 516
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 201
; LENGTH: 592
; TYPE: PRT
; ORGANISM: HUMAN
US-60-229-512-201

Query Match 31.4%; Score 480; DB 47; Length 592;
Best Local Similarity 33.7%; Pred. No. 2e-41;
Matches 112; Conservative 54; Mismatches 112; Indels 54; Gaps 8;
Qy 3 ESEATPSTP-----GEFSKYFEPHGVL-----PP--FCRG-----K 34
Db 265 EEGREQPYTPTTILSPCRGENSGRSLRNMELIQDTSRPPLEYKGVPLIKYFAEA 324
Qy 35 MEEIANFPVRPSDVWIVTPKSGTSLLOEVVYVLSQGGADPDDIGLMLNIDEQPLVLEYPOP 94
Db 325 LGPLOSFOARDDLLINTYPKSGTTWVSQILDMIYQGGDEKCNRAPYVRVPLEVNDP 384
Qy 95 ----GLDIKELTSRLIKSHLPYRFLPSDLHNGDSKVIYMARNPKDLVSYQPHRSLR 150
Db 385 GEPGSLTLDKTPPPRLIKSHLPALLPQTLDDQKVKVYVARNPKDVAVSYHPRMEK 444
Qy 151 TMSYRGTFQECRRFMNDKLGYSWFHVFQEFWEHRMDSNVLFKXYEDMHRDLVTMVQL 210
Db 445 AHPEPGTWDSELEKFMAGEVSYGSHYQHVQEWELSRTHPVLYLFYEDMKENPKREIQKI 504
Qy 211 ARFLGVSCDKAQLEALTEHC-----HQLVDOCCNABALP-VGRGRVGL 252
Db 505 LEFVGRSLPEETMDPMVQHTSFKEKKNPMNTYTTVPQELMDHSIS-----PPMRKGMAGD 560
Qy 253 WKDIFTVMNEKFDLVYKQKMGKCDLTFDFYL 284
Db 561 WKTFTVAQNERFDADYAEKMGAGCSLSRFE 592

RESULT 82
US-09-791-537-52308
; Sequence 52308, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055

; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 52308
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-52308
Query Match 31.3%; Score 479; DB 27; Length 295;
Best Local Similarity 36.6%; Pred. No. 9.2e-42;
Matches 102; Conservative 50; Mismatches 101; Indels 26; Gaps 4;
Qy 24 GVRLPFFCRGMEETANFPVRPSDVWIVTPKSGTSLLOEVVYVLSQGGADPDEIGLMNID 83
Db 17 GVPLIKYFAELGPLQSFQARPDDLLINTYPKSGTTWVSQILDMIYQGGDEKCNRAPIY 76
Qy 84 EQLPVLEYPOP-----GLDIIKELTSRLIKSHLPYRFLPSDLHNGDSKVIYMARNPKDIAV 139
Db 77 VRVPLEVNDPGEPSGLETLDKTPPPRLIKSHLPALLPQTLDDQKVKVYVARNPKDVA 136
Qy 140 VSYQPHRSLRMTSYRGTFQECRRFMNDKLGYSWFHVFQEFWEHRMDSNVLFKXYEDM 199
Db 137 VSYHFRHMEKAHPPEPGTWDSELEKFMAGEVSYGSHYQHVQEWELSRTHPVLYLFYEDM 196
Qy 200 HRDLVTMVQELARFLGVSCDKAQLEALTEHC-----HQLVDOCCNABEA 242
Db 197 KENPKREIQKILEFVGRSLPBEETVDMVVEHTSFKEKMKTPMTNTYTVRRELMDHSIS---- 253
Qy 243 LP-VGRGRVGLWKDIFTVMNEKFDLVYKQKMGKCDLTF 280
Db 254 -PPMRKGMAGDKTFTTVAQNERFDADYAEKMGAGCSLTF 291
RESULT 83
PCT-US01-08631-57913
; Sequence 57913, Application PC/TUS0108631
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-049
; CURRENT APPLICATION NUMBER: PCT/US01/08631
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 57913
; LENGTH: 1305
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (649)..(695)
; OTHER INFORMATION: Serine/threonine specific protein phosphatases proteins
; OTHER INFORMATION: domain identified by eMATRIX, accession number BL00125B, p-value=
; OTHER INFORMATION: 1.000e-40, raw score of 21.48
; NAME/KEY: DOMAIN
; LOCATION: (593)..(850)
; OTHER INFORMATION: Ser/Thr protein phosphatase domain identified by Pfam,
; OTHER INFORMATION: accession name Stposphatase, E-value=2.6e-196, Pfam score of 665
PCT-US01-08631-57913
Query Match 31.0%; Score 474; DB 1; Length 1305;
Best Local Similarity 33.5%; Pred. No. 2.7e-40;
Matches 110; Conservative 54; Mismatches 110; Indels 54; Gaps 8;
Qy 3 ESEATPSTP-----GEFSKYFEPHGVL-----PP--FCRG-----K 34
Db 265 EEGREQPYTPTTILSPCRGENSGRSLRNMELIQDTSRPPLEYKGVPLIKYFAEA 324
Qy 35 MEEIANFPVRPSDVWIVTPKSGTSLLOEVVYVLSQGGADPDEIGLMLNIDEQPLVLEYPOP 94

Db 325 LGPLOSFOARPDLLINTYPSKGTWVSQILDMIYQGGDLKCNRAPIYVRVPFLVNDP 384
Qy 95 ---GLDIIKELTSPLIKSHLPYRFLPSDLHNGDSKVIYARNPKDLVSVYQYHRSR 150
Db 385 GEPSGLETLKOTPPRLIKSHLPLALLPQTLLDQKVYVYARNPKDVAVSYYHFRMEK 444
Qy 151 TMSYRGTFQECRRRMNDKLGVSFWFHVQFWEHRMDSNVLFLEYEDMHRDLVTMVQOL 210
Db 445 AHPEPGTWDSELEKFMAGEVSGSWYQHVQEWELSRTHPVLYLFYEDMKENPKREIQKI 504
Qy 211 ARFLGVSCDKAQLEALTEHC-----HQLVDQCCNAEALP-VGRGRVGL 252
Db 505 LEFVGRSLPEETMDPMVQHTSKETKKNPMNTYTVPOELMDHSIS-----PFMRKGMAGD 560
Qy 253 WKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 561 WKTFTVAQNERFDADYAEKMGAGCSLSF 588
RESULT 84
US-10-450-763-57913
; Sequence 57913, Application US/10450763
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Cuetom
; SEQ ID NO 57913
; LENGTH: 1305
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: DOMAIN
; LOCATION: [649]..(695)
; OTHER INFORMATION: Serine/threonine specific protein phosphatases proteins
; OTHER INFORMATION: domain identified by eMATRIX, accession number BL001258, p-value=
; OTHER INFORMATION: 1.000e-40, raw score of 21.48
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (593)..(850)
; OTHER INFORMATION: Ser/Thr protein phosphatase domain identified by Pfam,
; OTHER INFORMATION: accession name Stposphatase, E-value=2.6e-196, Pfam score of 665
US-10-450-763-57913

Query Match 31.0%; Score 474; DB 34; Length 1305;
Best Local Similarity 33.5%; Pred. No. 2.7e-40;
Matches 110; Conservative 54; Mismatches 110; Indels 54; Gaps 8;
Qy 3 ESEATPSTP-----GEPSKYPEFHGVRL-----PP--FCRG-----K 34
Db 265 EEGREQVTPYTPILSPCRQGENSGSRELNMELIQDTSRPPLEYKVGVP LIKYFAEA 324
Qy 35 MEEIANFPVRSDWIVTPKSGTSLLOEVVYVLSQGDPPDGLMNTDEQLPVLEYQP 94
Db 325 LGPLOSFOARPDLLINTYPSKGTWVSQILDMIYQGGDLKCNRAPIYVRVPFLVNDP 384
Qy 95 ---GLDIIKELTSPLIKSHLPYRFLPSDLHNGDSKVIYARNPKDLVSVYQYHRSR 150
Db 385 GEPSGLETLKOTPPRLIKSHLPLALLPQTLLDQKVYVYARNPKDVAVSYYHFRMEK 444
Qy 151 TMSYRGTFQECRRRMNDKLGVSFWFHVQFWEHRMDSNVLFLEYEDMHRDLVTMVQOL 210
Db 445 AHPEPGTWDSELEKFMAGEVSGSWYQHVQEWELSRTHPVLYLFYEDMKENPKREIQKI 504

Qy 211 ARFLGVSCDKAQLEALTEHC-----HQLVDQCCNAEALP-VGRGRVGL 252
Db 505 LEFVGRSLPEETMDPMVQHTSKETKKNPMNTYTVPOELMDHSIS-----PFMRKGMAGD 560
Qy 253 WKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 561 WKTFTVAQNERFDADYAEKMGAGCSLSF 588
RESULT 85
PCT-US01-41986-17
; Sequence 17, Application PC/TUS0141986
; GENERAL INFORMATION:
; APPLICANT: Trustees of Boston University
; TITLE OF INVENTION: A NEURAL SPECIFIC CYTOSOLIC
; FILE REFERENCE: SULFOTRANSFERASE FOR DRUG SCREENING
; CURRENT APPLICATION NUMBER: PCT/US01/41986
; CURRENT FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: 60/229,929
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: DOMAIN
; LOCATION: [472]..(52e-41)
; OTHER INFORMATION: 30.8%; Score 472; DB 1; Length 295;
; OTHER INFORMATION: Best Local Similarity 35.8%; Pred. No. 5.2e-41;
; OTHER INFORMATION: Matches 100; Conservative 52; Mismatches 101; Indels 26; Gaps 4;
Qy 24 GVRLLPPFCRGKXMEIANPFPVRPSDVWIVTPKSGTSLLOEVVYVLSQGDPPDGLMNTDEQLPVLEYQP 83
Db 17 GVPLIKYPAEALGPLQSFQARPDLLINTYPSKGTWVSQILDMIYQGGDLKCNRAPIY 76
Qy 84 EQLPVLEYQP-----GLDIIKELTSPLIKSHLPYRFLPSDLHNGDSKVIYARNPKDLV 139
Db 77 VRVPFLVNDPGEPSGLETLLKTPPRLIKSHLPLALLPQTLLDQKVYVYARNPKDVA 136
Qy 140 VSYTFQHRSLRMTSYRGTFQECRRRMNDKLGVSFWFHVQFWEHRMDSNVLFLEYEDM 199
Db 137 VSYVHFHMERKHAPEPGTWDSELEKFMAGEVSGSWYQHVQEWELSRTHPVLYLFYEDM 196
Qy 200 HRDLVTMVQOLARFLGVSCDKAQLEALTEHC-----HQLVDQCCNAE 242
Db 197 KENPKREIQKILEFVGRSLPEETMDPMVQHTSKEMKKNPMNTYTVPOELMDHSIS--- 253
Qy 243 LP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 254 -PFMRKGMAGDWKTTFTVAQNERFDADYAEKMGAGCSLSF 291
RESULT 86
PCT-US05-01474-4
; Sequence 4, Application PC/TUS0501474
; GENERAL INFORMATION:
; APPLICANT: Squires, James E.
; APPLICANT: Lin, Zhihong E.
; TITLE OF INVENTION: Genetic Markers for Skatole Metabolism
; FILE REFERENCE: P06153U03
; CURRENT APPLICATION NUMBER: PCT/US05/01474
; CURRENT FILING DATE: 2005-01-18
; PRIOR APPLICATION NUMBER: 10/769,507
; PRIOR FILING DATE: 2004-01-30
; PRIOR APPLICATION NUMBER: 10/024,628
; PRIOR FILING DATE: 2001-11-23
; PRIOR APPLICATION NUMBER: 09/288,037
; PRIOR FILING DATE: 1999-04-08
; PRIOR APPLICATION NUMBER: 60/081,037
; PRIOR FILING DATE: 1998-04-08
; NUMBER OF SEQ ID NOS: 14

```

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-01474-4

Query Match      30.8%; Score 472; DB 1; Length 295;
Best Local Similarity 35.8%; Pred. No. 5.2e-41;
Matches 100; Conservative 52; Mismatches 101; Indels 26; Gaps 4;

QY 24 GVRLLPFCRGKMEETANFVRPSDVWIVTPKSGTSLLOEVVYLVSQGADPDEIGLMNID 83
DB 17 GVLPIKYFAEALGPLQSFQARPDLLINTYPKSGTTWVSQILDMIYQGGDLEKCNRAPIY 76

QY 84 EQLPVLEYPOP---GLDIIKELTSPLRIKSHLPYRFLPSDLHNGDSKVIYMARNPCKDLV 139
DB 77 VRVPFLEVNDPGEPSGLETLDKTPPRLIKSHLPLALLPOTLLDQKVKVYVARNPKDVA 136

QY 140 VSYIQFHRSLRTMSYRGTFQFECRRFMNDKLGYSWFEBHQVFEWHRMDSNVLFKYEDM 199
DB 137 VSYTHFRHMEKAHPPEGTWDSFLEKFMAGEVSYGSYQHWQEWELSRTHPVLYLFYEDM 196

QY 200 HRDLVTWVEQLARFLGVSCDKAQLAALTEHC-----HQLVDQCCNAEA 242
DB 197 KENPKREIQKILEFVGRSLPEETDMFVQHTSFKEMKKNPMTNVTYVPQELMDHSIS--- 253

QY 243 LP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
DB 254 -PFMRKGMAGDWKTTFTVAQNERFDADYAEKMGACCSLSF 291

RESULT 87
US-08-484-878-7
; Sequence 7, Application US/08484878
; GENERAL INFORMATION:
; APPLICANT: Beer-Romero, Peggy
; APPLICANT: Draetta, Giulio
; APPLICANT: Rolfe, Mark
; TITLE OF INVENTION: E6AP-Binding Proteins
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,878
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MII-044
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 295 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-484-878-7

Query Match      30.8%; Score 472; DB 14; Length 295;
Best Local Similarity 35.8%; Pred. No. 5.2e-41;
Matches 100; Conservative 52; Mismatches 101; Indels 26; Gaps 4;

QY 24 GVRLLPFCRGKMEETANFVRPSDVWIVTPKSGTSLLOEVVYLVSQGADPDEIGLMNID 83
DB 17 GVLPIKYFAEALGPLQSFQARPDLLINTYPKSGTTWVSQILDMIYQGGDLEKCNRAPIY 76

QY 84 EQLPVLEYPOP---GLDIIKELTSPLRIKSHLPYRFLPSDLHNGDSKVIYMARNPCKDLV 139
DB 77 VRVPFLEVNDPGEPSGLETLDKTPPRLIKSHLPLALLPOTLLDQKVKVYVARNPKDVA 136

QY 140 VSYIQFHRSLRTMSYRGTFQFECRRFMNDKLGYSWFEBHQVFEWHRMDSNVLFKYEDM 199
DB 137 VSYTHFRHMEKAHPPEGTWDSFLEKFMAGEVSYGSYQHWQEWELSRTHPVLYLFYEDM 196

QY 200 HRDLVTWVEQLARFLGVSCDKAQLAALTEHC-----HQLVDQCCNAEA 242
DB 197 KENPKREIQKILEFVGRSLPEETDMFVQHTSFKEMKKNPMTNVTYVPQELMDHSIS--- 253

QY 243 LP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
DB 254 -PFMRKGMAGDWKTTFTVAQNERFDADYAEKMGACCSLSF 291

RESULT 88
US-09-791-537-67010
; Sequence 67010, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 67010
; LENGTH: 295
; TYPE: PRT
; ORGANISM: pdb 1CJWA
US-09-791-537-67010

Query Match      30.8%; Score 472; DB 27; Length 295;
Best Local Similarity 35.8%; Pred. No. 5.2e-41;
Matches 100; Conservative 52; Mismatches 101; Indels 26; Gaps 4;

QY 24 GVRLLPFCRGKMEETANFVRPSDVWIVTPKSGTSLLOEVVYLVSQGADPDEIGLMNID 83
DB 17 GVLPIKYFAEALGPLQSFQARPDLLINTYPKSGTTWVSQILDMIYQGGDLEKCNRAPIY 76

QY 84 EQLPVLEYPOP---GLDIIKELTSPLRIKSHLPYRFLPSDLHNGDSKVIYMARNPCKDLV 139
DB 77 VRVPFLEVNDPGEPSGLETLDKTPPRLIKSHLPLALLPOTLLDQKVKVYVARNPKDVA 136

QY 140 VSYIQFHRSLRTMSYRGTFQFECRRFMNDKLGYSWFEBHQVFEWHRMDSNVLFKYEDM 199
DB 137 VSYTHFRHMEKAHPPEGTWDSFLEKFMAGEVSYGSYQHWQEWELSRTHPVLYLFYEDM 196

QY 200 HRDLVTWVEQLARFLGVSCDKAQLAALTEHC-----HQLVDQCCNAEA 242
DB 197 KENPKREIQKILEFVGRSLPEETDMFVQHTSFKEMKKNPMTNVTYVPQELMDHSIS--- 253

QY 243 LP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
DB 254 -PFMRKGMAGDWKTTFTVAQNERFDADYAEKMGACCSLSF 291

RESULT 89
US-10-170-205E-29903
; Sequence 29903, Application US/10170205E
; GENERAL INFORMATION:
; APPLICANT: ADAMS, Mark
```



```

Db      :      :      :      :      :      :      :      :      :      :
197 KENPREIOKILFEVGRSLPEETDMFQVHTSFEMKKKNPMTNTTVPQLMDHSIS---253

Qy      :      :      :      :      :      :      :      :      :      :
243 LP-VGGRGVGLWKDFTYSMNEKFPLVYKQMGKCDLTF 280

Db      :      :      :      :      :      :      :      :      :      :
254 -PMPRIKGAGDKWTTFTVAQNERFPDAEAKMGAGCSLSF 291

```

RESULT 99

```

US-10-170-205E-17237
; Sequence 17237, Application US/10170205E
; GENERAL INFORMATION:
; APPLICANT: ADAMS, Mark
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN
; FILE REFERENCE: CL001381
; CURRENT APPLICATION NUMBER: US/10/170,205E
; CURRENT FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 40312
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 17237
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-170-205E-17237

```

[illegible]

```

RESULT 100
US-09-791-537-112794
; Sequence 112794, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 112794
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-791-537-112794

```

Query Match	30.5%	Score 467;	DB 27;	Length 296;
Best Local Similarity	35.1%	Pred. No. 1.8e-40;		

[illegible]

Search completed: May 17, 2006, 11:16:34
Job time : 606 secs

THIS PAGE BLANK (USPTO)

ALIGNMENTS

RESULT 1
US-10-427-631-7

; Sequence 7, Application US/10427631
; Publication No. US20030175923A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION; TANG, Y. Tom;
; APPLICANT: CORLEY, Neil C.; GUEGLER, Karl J.;
; APPLICANT: BAUGHN, Mariah R.; LAL, Preeti G.;
; APPLICANT: YUE, Henry; HILLMAN, Jennifer L.;
; APPLICANT: AZINZAT, Yalda
; TITLE OF INVENTION: HUMAN TRANSFERASE PROTEINS
; FILE REFERENCE: PF-0592-1 DIV
; CURRENT APPLICATION NUMBER: US/10/427,631
; CURRENT FILING DATE: 2003-04-29
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: PCT/US99/20989
; PRIOR FILING DATE: 1999-09-09
; PRIOR APPLICATION NUMBER: US 60/172,220
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: US 60/155,248
; PRIOR FILING DATE: 1998-11-04
; PRIOR APPLICATION NUMBER: US 60/133,642
; PRIOR FILING DATE: 1999-05-11
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PERL Program
; SEQ ID NO 7
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; FEATURE:
; OTHER INFORMATION: Incyte ID No. US20030175923A1 1784742CD1
US-10-427-631-7

Query Match 100.0%; Score 1530; DB 4; Length 284;
Best Local Similarity 100.0%; Pred. No. 1e-144;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MAESEAEPTPTGPFESKYFEGHGVRLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60
Db 1 MAESEAEPTPTGPFESKYFEGHGVRLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60
Qy 61 LQEVVYLVSQGADPDEIGLMNIDQLPVLEYPOPLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LQEVVYLVSQGADPDEIGLMNIDQLPVLEYPOPLDIIKELTSPRLIKSHLPYRFLPSD 120
Qy 121 LHNGDSKVIYARNPKDLVSVYQFHRSLRTWSYRGTFQFCRRPMNDKLGYSWFHVQ 180
Db 121 LHNGDSKVIYARNPKDLVSVYQFHRSLRTWSYRGTFQFCRRPMNDKLGYSWFHVQ 180
Qy 181 EFWEHRMDSNVLFKLYEDMHRDLVTWVEQLARFLGVSCDKAQLAEALTECHQLVDQCCNA 240
Db 181 EFWEHRMDSNVLFKLYEDMHRDLVTWVEQLARFLGVSCDKAQLAEALTECHQLVDQCCNA 240
Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284

RESULT 2

US-10-757-262-110
; Sequence 110, Application US/10757262
; Publication No. US20040197825A1
; GENERAL INFORMATION:
; APPLICANT: Karicheti, Venkateswarlu
; APPLICANT: Silos-Santiago, Inmaculada
; APPLICANT: Eliasof, Scott D.

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: UROLOGICAL DISORDERS USING 44390, 54181, 211, 5687, 884,
; TITLE OF INVENTION: 1405, 636, 4421, 5410, 30905, 2045, 16405, 18560, 2047,
; TITLE OF INVENTION: 33751, 52872, 14063, 20739, 32544, 43239, 44373, 51164,
; TITLE OF INVENTION: 53010, 16852, 1587, 2207, 22245, 2387, 52908, 69112, 14990,
; TITLE OF INVENTION: 18547, 115, 579, 15985, 15625, 760, 18603, 2395, 2554, 8675,
; TITLE OF INVENTION: 32720, 4809, 14303, 16816, 17827, 32620, 577, 619, 1423,
; TITLE OF INVENTION: 2158, 8263, 15402, 16209, 16386, 21165, 30911, 41897, 1643,
; TITLE OF INVENTION: 2543, 9626, 13231, 32409, 84260, 2882, 8203, 32678 OR
; TITLE OF INVENTION: 55053
; FILE REFERENCE: MPI03-007P1RNMNIM
; CURRENT APPLICATION NUMBER: US/10/757,262
; CURRENT FILING DATE: 2004-01-14
; PRIOR APPLICATION NUMBER: US 60/440,318
; PRIOR FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: US 60/444,783
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: US 60/457,901
; PRIOR FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: US 60/468,775
; PRIOR FILING DATE: 2003-05-08
; PRIOR APPLICATION NUMBER: US 60/471,614
; PRIOR FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: US 60/478,742
; PRIOR FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: US 60/488,529
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/491,156
; PRIOR FILING DATE: 2003-07-30
; PRIOR APPLICATION NUMBER: US 60/499,594
; PRIOR FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US 60/506,332
; PRIOR FILING DATE: 2003-09-26
; NUMBER OF SEQ ID NOS: 136
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 110
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-757-262-110
Query Match 100.0%; Score 1530; DB 4; Length 284;
Best Local Similarity 100.0%; Pred. No. 1e-144;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MAESEAEPTPTGPFESKYFEGHGVRLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60
Db 1 MAESEAEPTPTGPFESKYFEGHGVRLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60
Qy 61 LQEVVYLVSQGADPDEIGLMNIDQLPVLEYPOPLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LQEVVYLVSQGADPDEIGLMNIDQLPVLEYPOPLDIIKELTSPRLIKSHLPYRFLPSD 120
Qy 121 LHNGDSKVIYARNPKDLVSVYQFHRSLRTWSYRGTFQFCRRPMNDKLGYSWFHVQ 180
Db 121 LHNGDSKVIYARNPKDLVSVYQFHRSLRTWSYRGTFQFCRRPMNDKLGYSWFHVQ 180
Qy 181 EFWEHRMDSNVLFKLYEDMHRDLVTWVEQLARFLGVSCDKAQLAEALTECHQLVDQCCNA 240
Db 181 EFWEHRMDSNVLFKLYEDMHRDLVTWVEQLARFLGVSCDKAQLAEALTECHQLVDQCCNA 240
Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284

RESULT 3

US-10-768-158-2
; Sequence 2, Application US/10768158
; Publication No. US20040204359A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Silos-Santiago, Inmaculada

```
; APPLICANT: Karicheti, Venkateswarlu
; APPLICANT: Eliasof, Scott D.
; TITLE OF INVENTION: PAIN AND PAINFUL DISORDERS USING 16386, 15402, 21165, 1423,
; TITLE OF INVENTION: 636, 12303, 21425, 27410, 38554, 38555, 55063, 57145, 59914,
; TITLE OF INVENTION: 94921, 16852, 33260, 58573, 30911, 85913, 14303, 16816,
; TITLE OF INVENTION: 17827 OR 32620
; FILE REFERENCE: MPI03-012P1RNM1M
; CURRENT APPLICATION NUMBER: US/10/768,158
; CURRENT FILING DATE: 2004-01-30
; PRIOR FILING DATE: 2004-01-30
; PRIOR FILING DATE: 2003-02-04
; PRIOR FILING DATE: 2003-02-04
; PRIOR FILING DATE: 2003-03-05
; PRIOR FILING DATE: 2003-03-13
; PRIOR FILING DATE: 2003-03-13
; PRIOR FILING DATE: 2003-06-16
; PRIOR FILING DATE: 2003-07-30
; PRIOR FILING DATE: 2003-07-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo Sapiens
; US-10-768-158-2

Query Match      100.0%; Score 1530; DB 4; Length 284;
Best Local Similarity 100.0%; Pred. No. 1e-144;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAESEAETPTPGEPESKYFFHGVRLPPFCRGKMEIEANFPVRPSDVWIVTPKSGTSL 60
Db 1 MAESEAETPTPGEPESKYFFHGVRLPPFCRGKMEIEANFPVRPSDVWIVTPKSGTSL 60

Qy 61 LQEVVYLVVSQADPDEIGLMNIDEQLPVLEYPOQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LQEVVYLVVSQADPDEIGLMNIDEQLPVLEYPOQGLDIIKELTSPRLIKSHLPYRFLPSD 120

Qy 121 LHNGDSKVIYMARNP KD LVSYQPHRSRLTMSYRGTFQFCRRRPMNDKLGYSWFHVQ 180
Db 121 LHNGDSKVIYMARNP KD LVSYQPHRSRLTMSYRGTFQFCRRRPMNDKLGYSWFHVQ 180

Qy 181 EFWEHRMDSNVLFKYEYEDHMDRLVTWVEQLARFLGVS CDKAQLEALTEHCHQLVDOCCNA 240
Db 181 EFWEHRMDSNVLFKYEYEDHMDRLVTWVEQLARFLGVS CDKAQLEALTEHCHQLVDOCCNA 240

Qy 241 EALPVGRGRVGLWKDIFTVSNNEKFDLVYKQMGKCDLTFDFYL 284
Db 241 EALPVGRGRVGLWKDIFTVSNNEKFDLVYKQMGKCDLTFDFYL 284

RESULT 4
US-10-205-331-46
; Sequence 46, Application US/10205331
; Publication No. US20040058326A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; APPLICANT: Lee, Kevin
; APPLICANT: Dixon, Allstair
; APPLICANT: Brooksbank, Robert
; APPLICANT: Pinnoch, Robert
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
; FILE REFERENCE: WL-A-018199
; CURRENT APPLICATION NUMBER: US/10/205,331
; CURRENT FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: GB 0118354.0
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 117
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 46
; LENGTH: 284
```

```
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Sulfotransferase-like protein
US-10-205-331-46

Query Match      98.4%; Score 1506; DB 4; Length 284;
Best Local Similarity 97.9%; Pred. No. 2.6e-142;
Matches 278; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MAESEAETPTPGEPESKYFFHGVRLPPFCRGKMEIEANFPVRPSDVWIVTPKSGTSL 60
Db 1 MAESEAETPTPGEPESKYFFHGVRLPPFCRGKMEIEANFPVRPSDVWIVTPKSGTSL 60

Qy 61 LQEVVYLVVSQADPDEIGLMNIDEQLPVLEYPOQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LQEVVYLVVSQADPDEIGLMNIDEQLPVLEYPOQGLDIIKELTSPRLIKSHLPYRFLPSD 120

Qy 121 LHNGDSKVIYMARNP KD LVSYQPHRSRLTMSYRGTFQFCRRRPMNDKLGYSWFHVQ 180
Db 121 LHNGDSKVIYMARNP KD LVSYQPHRSRLTMSYRGTFQFCRRRPMNDKLGYSWFHVQ 180

Qy 181 EFWEHRMDSNVLFKYEYEDHMDRLVTWVEQLARFLGVS CDKAQLEALTEHCHQLVDOCCNA 240
Db 181 EFWEHRMDSNVLFKYEYEDHMDRLVTWVEQLARFLGVS CDKAQLEALTEHCHQLVDOCCNA 240

Qy 241 EALPVGRGRVGLWKDIFTVSNNEKFDLVYKQMGKCDLTFDFYL 284
Db 241 EALPVGRGRVGLWKDIFTVSNNEKFDLVYKQMGKCDLTFDFYL 284

RESULT 5
US-10-072-012-667
; Sequence 667, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
```

```
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 667
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-072-012-667

Query Match      31.4%; Score 480; DB 4; Length 296;
Best Local Similarity 35.6%; Pred. No. 2.7e-39;
Matches 103; Conservative 57; Mismatches 107; Indels 22; Gaps 5;

Qy 16 ESKYFERHGVRLPPFCRCGMKEIANFPVRPSDVMTVTYPKSGTLLQEVVYLVSGADPD 75
Db 10 QTKUKEVAGIFLQAPTVDNWKQIQTFEAKPDLLICTTPKSGTTWQIWDIMIEQNGDVE 69

Qy 76 EIGLNMIDEQLVLEY---POP-GLDIIKELTSPRLIKSHLPYRFLPSPDLHNGDSKVLYM 131
Db 70 KCRRTIIQHRHPFIEWARPPQSGVDKANEMPAPRILRTHLPTQLLPSPFWTNCKFLYV 129

Qy 132 ARNPKDLVVSYYQFHRSLRTHSYRGTFOFCRFRFNDKLGYSWFHVEHWRMDSNV 191
Db 130 ARNAKDCMVSYHYFHYRMSQVLPEPQTWDEYFETFTNGKVSFWDHVKWMEIRDKYQI 189

Qy 192 LFLKYEDMHRDLVTMVEQLARFLGVS-----CDKAQLEALTEHCHQLVDQCNAEALP- - 244
Db 190 LFLFYEDMKRNPKEHIQKVMQFMGKNLDEVDVVKLVLTSEPKMKE--NPMTNSTAPKS 247

Qy 245 -----VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL 284
Db 248 ILDQISPFMRKGTVDGKNHFTVAQNERFDEIYKQKMGRTSLNFSMEL 296

RESULT 6
US-10-450-763-57913
; Sequence 57913, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 57913
; LENGTH: 1305
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: {6491}..(695)
; OTHER INFORMATION: Serine/threonine specific protein phosphatases
; OTHER INFORMATION: domain identified by eMATRIX, accession number BL00125B, p-value=
; OTHER INFORMATION: 1.000e-40, raw score of 21.48
; FEATURE:
```

```
; NAME/KEY: DOMAIN
; LOCATION: (593)..(850)
; OTHER INFORMATION: Ser/Thr protein phosphatase domain identified by Pfam,
; OTHER INFORMATION: accession name Stposphatase, E-value=2.6e-196, Pfam score of 665
US-10-450-763-57913

Query Match      31.0%; Score 474; DB 5; Length 1305;
Best Local Similarity 33.5%; Pred. No. 8.6e-38;
Matches 110; Conservative 54; Mismatches 110; Indels 54; Gaps 8;

Qy 3 ESEAEETSTP-----GEPEKYFEFPHGVRL-----PP--FCRG-----K 34
Db 265 EEGREQPTPTTTPILSPCRQGENRSGSRELNNMELIQDTSRPPELVYKGVPLIKYPAEA 324

Qy 35 MEEIANFPVRPSDVMTVTYPKSGTLLQEVVYLVSGADPDDEIGLMDIDQLPVLEYPQP 94
Db 325 LGPLQSFQARPDLLINTYPKSGTTWVSQILDMTYQGGDLKCNRAPIYVRVPFLEVNDP 384

Qy 95 ---GLDIIKELTSPRLIKSHLPYRFLPSPDLHNGDSKVVIYMARNPDKLVSYYYQFHRSLR 150
Db 385 GEPSGLTLDKTPPPRLIKSHLPALLPQTLDDQKVYVYVARNPKDVANVSYHFRHMEK 444

Qy 151 TMSYRGTFQFCRFRFNDKLGYSWFHVEHWRMDSNVLFKYEDMHRDLVTMVEQL 210
Db 445 AHPEPGTWSDFLEXFMAGEVSYSYQHWQEWELSRTHPVLYLFYEDMKENPKREIQKI 504

Qy 211 ARFLGVSCDKAQLEALTEHC-----HQLVDQCNAEALP-VGRGRVGL 252
Db 505 LEFVGRSLPEETDMFVQHTSFKETKNPMNTYTPQELMDHSIS----PFMRKMGAGD 560

Qy 253 WKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 561 WKTFTVAQNERFDADVAEKMGACSLSF 588

RESULT 7
US-09-854-122-21
; Sequence 21, Application US/09854122
; Patent No. US20020016980A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTE, RANDALL S.
; APPLICANT: SMITH, ROBERT
; TITLE OF INVENTION: TRANSGENIC PLANTS INCORPORATING TRAIT OF ZOSTERA MARINA
; FILE REFERENCE: PHA-007.01
; CURRENT APPLICATION NUMBER: US/09/854,122
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/202,529
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-854-122-21

Query Match      30.8%; Score 472; DB 3; Length 295;
Best Local Similarity 35.8%; Pred. No. 1.7e-38;
Matches 100; Conservative 52; Mismatches 101; Indels 26; Gaps 4;

Qy 24 GVRLPFFCRGKWEIANKFPVRPSDVMTVTYPKSGTLLQEVVYLVSGADPDDEIGLNMID 83
Db 17 GVPLIKYPAEALGFLQSFQARPDLLINTYPKSGTTWVSQILDMTYQGGDLKCNRAPIY 76

Qy 84 EQLPVLEYQPQ---GLDIIKELTSPRLIKSHLPYRFLPSPDLHNGDSKVVIYMARNPDKLV 139
Db 77 VRVPFLEVNDPGEPSGLETLDKTPPRLIKSHLPALLPQTLDDQKVYVYVARNPKDVA 136

Qy 140 VSYIQFHRSLRTHSYRGTFOFCRFRFNDKLGYSWFHVEHWRMDSNVLFKYEDM 199
Db 137 VSYHFRHMEKAPPEPGTWSDFLEKFMAGEVSYSYQHWQEWELSRTHPVLYLFYEDM 196

Qy 200 HRDLVTMVEQLARFLGVSCDKAQLEALTEHC-----HQLVDQCNAE 242
```


Db 197 KENPKREIQKILEFVGRSLPEETDMFVQHTSFKEMKKNPMTNTYTPQELMDHSIS--- 253
Qy 243 LP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 254 -PFMRKGMAGDWKTTFTTVAQNERFDADYAEKMGAGCSLSF 291

RESULT 8

US-10-769-507-4
; Sequence 4, Application US/10769507
; Publication No. US20050019788A1
; GENERAL INFORMATION:
; APPLICANT: Squires, James E.
; TITLE OF INVENTION: Genetic Markers for Skatole Metabolism
; FILE REFERENCE: P06153U03
; CURRENT APPLICATION NUMBER: US/10769,507
; CURRENT FILING DATE: 2004-01-30
; PRIOR APPLICATION NUMBER: 10/024,628
; PRIOR FILING DATE: 2001-11-23
; PRIOR APPLICATION NUMBER: 09/288,037
; PRIOR FILING DATE: 1999-04-08
; PRIOR APPLICATION NUMBER: 60/081,037
; PRIOR FILING DATE: 1998-04-08
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-769-507-4

Query Match 30.8%; Score 472; DB 5; Length 295;
Best Local Similarity 35.8%; Pred. No. 1.7e-38;
Matches 100; Conservative 52; Mismatches 101; Indels 26; Gaps 4;
Qy 24 GVRLLPPFCRGKMEIANTFVRPSDWIVTYPKSGTSLLOEVVYLVSQAGDDEIGLMNID 83
Db 17 GVPLIKYFAEALGPLQSFQARPDLLINTYPKSGTTWVSQILDMIYQGGDLEKCNRAPIY 76
Qy 84 EQLPVLEYPQP---GLDIIKELTSPRLIKSHLPVRFPLPSDLHNGDSKVIYARNPKDLV 139
Db 77 VRVFLVNDPGEPSGLETLDTPPRLIKSHLPALLPQTLLOKVKVYVARNPKDVA 136
Qy 140 VSYTQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVFWEHRMDSNVFLFKYEDM 199
Db 137 VSYHFRHMEKAHPPEPGTWSDFLEKFMAGEVSYSGWYQHVQEWELSRTHPVLVLYFYEDM 196
Qy 200 HRDLVTWVEQLARFLGVSCDKAQLAETHC-----HQLVDQCCNAEA 242
Db 197 KENPKREIQKILEFVGRSLPEETDMFVQHTSFKEMKKNPMTNTYTPQELMDHSIS--- 253
Qy 243 LP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 254 -PFMRKGMAGDWKTTFTTVAQNERFDADYAEKMGAGCSLSF 291

RESULT 9

US-11-033-030-21
; Sequence 21, Application US/11033030
; Publication No. US20060053510A1
; GENERAL INFORMATION:
; APPLICANT: SMITH, ROBERT
; APPLICANT: ALBERTE, RANDALL S.
; TITLE OF INVENTION: TRANSGENIC PLANTS INCORPORATING TRAIT OF ZOSTERA MARINA
; FILE REFERENCE: PHA-007.01
; CURRENT APPLICATION NUMBER: US/11/033,030
; CURRENT FILING DATE: 2005-01-11
; PRIOR APPLICATION NUMBER: US/09/854,122
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/202,529
; PRIOR FILING DATE: 2000-05-10

; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 21
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-033-030-21

Query Match 30.8%; Score 472; DB 6; Length 295;
Best Local Similarity 35.8%; Pred. No. 1.7e-38;
Matches 100; Conservative 52; Mismatches 101; Indels 26; Gaps 4;
Qy 24 GVRLLPPFCRGKMEIANTFVRPSDWIVTYPKSGTSLLOEVVYLVSQAGDDEIGLMNID 83
Db 17 GVPLIKYFAEALGPLQSFQARPDLLINTYPKSGTTWVSQILDMIYQGGDLEKCNRAPIY 76
Qy 84 EQLPVLEYPQP---GLDIIKELTSPRLIKSHLPVRFPLPSDLHNGDSKVIYARNPKDLV 139
Db 77 VRVFLVNDPGEPSGLETLDTPPRLIKSHLPALLPQTLLOKVKVYVARNPKDVA 136
Qy 140 VSYTQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVFWEHRMDSNVFLFKYEDM 199
Db 137 VSYHFRHMEKAHPPEPGTWSDFLEKFMAGEVSYSGWYQHVQEWELSRTHPVLVLYFYEDM 196
Qy 200 HRDLVTWVEQLARFLGVSCDKAQLAETHC-----HQLVDQCCNAEA 242
Db 197 KENPKREIQKILEFVGRSLPEETDMFVQHTSFKEMKKNPMTNTYTPQELMDHSIS--- 253
Qy 243 LP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 254 -PFMRKGMAGDWKTTFTTVAQNERFDADYAEKMGAGCSLSF 291

RESULT 10

US-10-072-012-666
; Sequence 666, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31


```
; Publication No. US20050019788A1
; GENERAL INFORMATION:
; APPLICANT: Squires, James E.
; TITLE OF INVENTION: Genetic Markers for Skatole Metabolism
; FILE REFERENCE: P06153US03
; CURRENT APPLICATION NUMBER: US/10/769,507
; PRIOR FILING DATE: 2004-01-30
; PRIOR APPLICATION NUMBER: 10/024,628
; PRIOR FILING DATE: 2001-11-23
; PRIOR APPLICATION NUMBER: 09/288,037
; PRIOR FILING DATE: 1999-04-08
; PRIOR APPLICATION NUMBER: 60/081,037
; PRIOR FILING DATE: 1998-04-08
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-769-507-3

Query Match      29.8%; Score 456; DB 5; Length 295;
Best Local Similarity 33.7%; Pred. No. 7e-37;
Matches 94; Conservative 58; Mismatches 101; Indels 26; Gaps 4;

QY 24 GVRLPFCRGKMEIEANFPVRPSDVWIVTPKSGTSLLOEVVYLVSQGADPDEIGLMNID 83
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 17 GVPLIKYFAEALGPLSQFQARPDLLISTYPKSGTTWVSQILDMIYQGGDLKCHRAPIF 76

QY 84 EQLPVLEYPOP-----GLDIIKELTSRLIKSHLPYRFLPSDLHNGDSKVIYMARPKDLV 139
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 77 MRVPFELEFKIPGIPSGMETLKNTPAPRLKTLHLPTLLPQTLDDQKVYVYVARNAKOVA 136

QY 140 VSYIQFHRSLRTMSYRGTFQFCRRFMDNDKLGYSWFEHVOEFWEHRMDSNVFLFKYEDM 199
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
Db 137 VSYHYFYHMAKVYPHPGTWDSFLEDFMAGEVSYGSWYQHVQEWELRTHPVLVLYFYEDM 196

QY 200 HRDLVTMVEQLARFLGVSCKAQLAELTEHC-----HQLVDOCCNAEALP-----V 245
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 197 KENPKREIQKILEFVGRSLPEETVEDIVQHTSFQEMKNAMTNYRTLPDDLHDHSISAFM 256

QY 246 GRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTF 280
   : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
Db 257 RKGITGDKWKTFTTVAQNERFEADYAEKMGAGCNLRF 291

RESULT 15
US-10-287-436A-372
; Sequence 372, Application US/10287436A
; Publication No. US2005020421A1
; GENERAL INFORMATION:
; APPLICANT: CHILDREN'S HOSPITAL MEDICAL CENTER
; TITLE OF INVENTION: METHOD FOR DIAGNOSIS AND TREATMENT OF
; FILE REFERENCE: 10872.514696
; CURRENT APPLICATION NUMBER: US/10/287,436A
; CURRENT FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: US 60/336,220
; PRIOR FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 1446
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 372
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-287-436A-372

Query Match      29.8%; Score 456; DB 5; Length 295;
Best Local Similarity 33.7%; Pred. No. 7e-37;
Matches 94; Conservative 58; Mismatches 101; Indels 26; Gaps 4;

QY 24 GVRLPFCRGKMEIEANFPVRPSDVWIVTPKSGTSLLOEVVYLVSQGADPDEIGLMNID 83
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 17 GVPLIKYFAEALGPLSQFQARPDLLISTYPKSGTTWVSQILDMIYQGGDLKCHRAPIF 76

QY 84 EQLPVLEYPOP-----GLDIIKELTSRLIKSHLPYRFLPSDLHNGDSKVIYMARPKDLV 139
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 77 MRVPFELEFKIPGIPSGMETLKNTPAPRLKTLHLPTLLPQTLDDQKVYVYVARNAKOVA 136

QY 140 VSYIQFHRSLRTMSYRGTFQFCRRFMDNDKLGYSWFEHVOEFWEHRMDSNVFLFKYEDM 199
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
Db 137 VSYHYFYHMAKVYPHPGTWDSFLEDFMAGEVSYGSWYQHVQEWELRTHPVLVLYFYEDM 196

QY 200 HRDLVTMVEQLARFLGVSCKAQLAELTEHC-----HQLVDOCCNAEALP-----HQLVDOCCNAE 242
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 197 KENPKREIQKILEFVGRSLPEETVDMVEHTSFQEMKNAMTNYRTVTRREFMDHSIS--- 253

QY 243 LP-VGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTF 280
   : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
Db 254 -PFMRKMGAGDKWKTFTTVAQNERFDADYAEKMGAGCSLSF 291

RESULT 16
US-10-287-436A-1246
; Sequence 1246, Application US/10287436A
```

```
; Publication No. US20050019788A1
; GENERAL INFORMATION:
; APPLICANT: Squires, James E.
; TITLE OF INVENTION: Genetic Markers for Skatole Metabolism
; FILE REFERENCE: P06153US03
; CURRENT APPLICATION NUMBER: US/10/769,507
; PRIOR FILING DATE: 2004-01-30
; PRIOR APPLICATION NUMBER: 10/024,628
; PRIOR FILING DATE: 2001-11-23
; PRIOR APPLICATION NUMBER: 09/288,037
; PRIOR FILING DATE: 1999-04-08
; PRIOR APPLICATION NUMBER: 60/081,037
; PRIOR FILING DATE: 1998-04-08
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 295
; TYPE: PRT
; ORGANISM: porcine
US-10-769-507-6

Query Match      29.8%; Score 456; DB 5; Length 295;
Best Local Similarity 33.7%; Pred. No. 7e-37;
Matches 94; Conservative 58; Mismatches 101; Indels 26; Gaps 4;

QY 24 GVRLPFCRGKMEIEANFPVRPSDVWIVTPKSGTSLLOEVVYLVSQGADPDEIGLMNID 83
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 17 GVPLIKYFAEALGPLSQFQARPDLLISTYPKSGTTWVSQILDMIYQGGDLKCHRAPIF 76

QY 84 EQLPVLEYPOP-----GLDIIKELTSRLIKSHLPYRFLPSDLHNGDSKVIYMARPKDLV 139
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 77 MRVPFELEFKIPGIPSGMETLKNTPAPRLKTLHLPTLLPQTLDDQKVYVYVARNAKOVA 136

QY 140 VSYIQFHRSLRTMSYRGTFQFCRRFMDNDKLGYSWFEHVOEFWEHRMDSNVFLFKYEDM 199
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
Db 137 VSYHYFYHMAKVYPHPGTWDSFLEDFMAGEVSYGSWYQHVQEWELRTHPVLVLYFYEDM 196

QY 200 HRDLVTMVEQLARFLGVSCKAQLAELTEHC-----HQLVDOCCNAEALP-----HQLVDOCCNAE 242
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 197 KENPKREIQKILEFVGRSLPEETVDMVEHTSFQEMKNAMTNYRTVTRREFMDHSIS--- 253

QY 243 LP-VGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTF 280
   : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
Db 254 -PFMRKMGAGDKWKTFTTVAQNERFDADYAEKMGAGCSLSF 291

RESULT 14
US-10-769-507-6
; Sequence 6, Application US/10769507
; Publication No. US20050019788A1
; GENERAL INFORMATION:
; APPLICANT: Squires, James E.
; TITLE OF INVENTION: Genetic Markers for Skatole Metabolism
; FILE REFERENCE: P06153US03
; CURRENT APPLICATION NUMBER: US/10/769,507
; CURRENT FILING DATE: 2004-01-30
; PRIOR APPLICATION NUMBER: 10/024,628
; PRIOR FILING DATE: 2001-11-23
; PRIOR APPLICATION NUMBER: 09/288,037
; PRIOR FILING DATE: 1999-04-08
; PRIOR APPLICATION NUMBER: 60/081,037
; PRIOR FILING DATE: 1998-04-08
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 295
; TYPE: PRT
; ORGANISM: porcine
US-10-769-507-6

Query Match      29.8%; Score 456; DB 5; Length 295;
```

```
; Publication No. US20050202421A1
; GENERAL INFORMATION:
; APPLICANT: CHILDREN'S HOSPITAL MEDICAL CENTER
; TITLE OF INVENTION: METHOD FOR DIAGNOSIS AND TREATMENT OF
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS
; FILE REFERENCE: 10872.514696
; CURRENT APPLICATION NUMBER: US/10/287,436A
; CURRENT FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: US 60/336,220
; PRIOR FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 1446
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1246
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-287-436A-1246

Query Match      29.8%; Score 456; DB 5; Length 295;
Best Local Similarity 33.7%; Pred. No. 7e-37;
Matches 94; Conservative 58; Mismatches 101; Indels 26; Gaps 4;

QY 24 GVRLLPFCRCGRKMEIEANFPVRPSDVWIVTPKSGTSLQEVVYLVVSQGDADPDEIGLMNID 83
Db 17 GVPLIKYFAEALGPLQSFQARPDDLLISTYPKSGTTWVSQILDMIYQGGDLEKCHRAPIF 76
QY 84 EQLPVLEYPOP-GLDIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNP KD LV 139
Db 77 MRVPFLEFKAPGPSGMEETLKDTAPRLLKTHLPLALLPQTLLDQKVKVYVARNAKDVA 136
QY 140 VSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWEHVOEFWEHRMDSNVFLFKYEDM 199
Db 137 VSYHFFHMAKVHPHPGPTWDSFLEKFMVGEVSYGSYQHVQEWELSRTHPVLVLYFYEDM 196
QY 200 HRDLVTMVEQLARFLGVSCDKAQLAEALTEHC-----HQLVDQCCNAEALP- 245
Db 197 KENPKREIQKILEFVGHSLPEETVDFMVQHTSFKEKMKKNPMTNTTVPQEFMDHSISPFM 256
QY 246 GRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 257 RKG MAGDWKTTTVAQNERFDADYAEK MAGCSLSF 291

RESULT 18
US-10-734-049A-245
; Sequence 245, Application US/10734049A
; Publication No. US20050042624A1
; GENERAL INFORMATION:
; APPLICANT: ITOH, Kyogo
; TITLE OF INVENTION: TUMOR ANTIGEN
; FILE REFERENCE: Q-78382
; CURRENT APPLICATION NUMBER: US/10/734,049A
; CURRENT FILING DATE: 2003-12-12
; PRIOR APPLICATION NUMBER: PCT/JP02/05799
; PRIOR FILING DATE: 2002-06-11
; PRIOR APPLICATION NUMBER: JP 2001/177058
; PRIOR FILING DATE: 2001-06-12
; PRIOR APPLICATION NUMBER: JP 2001/250728
; PRIOR FILING DATE: 2001-08-21
; NUMBER OF SEQ ID NOS: 408
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 245
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-734-049A-245

Query Match      29.5%; Score 451; DB 5; Length 295;
Best Local Similarity 33.8%; Pred. No. 2.2e-36;
Matches 93; Conservative 57; Mismatches 107; Indels 18; Gaps 3;

QY 24 GVRLLPFCRCGRKMEIEANFPVRPSDVWIVTPKSGTSLQEVVYLVVSQGDADPDEIGLMNID 83
Db 17 GVPLIKYFAEALGPLQSFQARPDDLLISTYPKSGTTWVSQILDMIYQGGDLEKCHRAPIF 76
QY 84 EQLPVLEYPOP-GLDIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNP KD LV 139
Db 77 MRVPFLEFKAPGPSGMEETLKDTAPRLLKTHLPLALLPQTLLDQKVKVYVARNAKDVA 136
QY 140 VSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWEHVOEFWEHRMDSNVFLFKYEDM 199
Db 137 VSYHFFHMAKVHPHPGPTWDSFLEKFMVGEVSYGSYQHVQEWELSRTHPVLVLYFYEDM 196
QY 200 HRDLVTMVEQLARFLGVSCDKAQLAEALTEHC-----HQLVDQCCNAEALP- 245
Db 197 KENPKREIQKILEFVGHSLPEETVDFMVQHTSFKEKMKKNPMTNTTVPQEFMDHSISPFM 256
QY 246 GRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 257 RKG MAGDWKTTTVAQNERFDADYAEK MAGCSLSF 291

RESULT 19
US-10-072-012-586
; Sequence 586, Application US/10072012
```

```

; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grobse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 586
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-072-012-586

Query Match      29.3%; Score 449; DB 4; Length 302;
Best Local Similarity 36.2%; Pred. No. 3 6e-36;
Matches 96; Conservative 52; Mismatches 99; Indels 18; Gaps 3;

QY 36 BEIANFVRPSDVMTVTPKSGTSLQEVVYLVSGADPDEIGLMNIDEQLPVLEYPOP- 94
Db 36 DKINWFQAKPDDLIIISTYPKAGTTWTQEIIVELIQNEGVDEKSKRAPHQRPFFLEMKIPS 95
QY 95 ---GLDIIKELTSPLIKSHLPYRPLPSDLHNGDSKVLYMARNPKDLVSYVYQFHSRLT 151
Db 96 LGSGLQEAHAMPSPRIILKTHLPFLLPSPLEKNCKIIYVARNPKNMVSYYHFORMNKA 155
QY 152 MSYRGTFQFCRRFMNDKLGYSWFHVEHWMDSNVLFKYEDMHRDLVTMVEOLA 211
Db 156 LPAPGTWEEYFETFLAGVCWGSWHEHVKGWEAKDKHRLVLYFYEDMKKPKHIEIQKLA 215
QY 212 RFLGVSCDKAQLEALTEHCHQLV---DQCCNAEALP-----VGRGRVGLWKDIF 257
Db 216 EFIGKLLDKVLDKIVHYTSFDMKQNPMPANYSSIPAEIMDHSISPPFMRKGAVGDKKHIF 275
QY 258 TVSMNEKFDLVYKQMGKCDLTFDF 282
Db 276 TVAQNERFDEYKXKMTDTRLTFHF 300

RESULT 21
US-10-072-012-668
; Sequence 668, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha

QY 212 RFLGVSCDKAQLEALTEHCHQLV---DQCCNAEALP-----VGRGRVGLWKDIF 257
Db 216 EFIGKLLDKVLDKIVHYTSFDMKQNPMPANYSSIPAEIMDHSISPPFMRKGAVGDKKHIF 275
QY 258 TVSMNEKFDLVYKQMGKCDLTFDF 282
Db 276 TVAQNERFDEYKXKMTDTRLTFHF 300

US-10-370-715B-590
; Sequence 590, Application US/10370715B
; Publication No. US20040258678A1
; GENERAL INFORMATION:
; Patin Docket Preview
; APPLICANT: BODARY, SARAH C.
; APPLICANT: CLARK, HILLARY
; APPLICANT: BRISDELL, HUNTE
; APPLICANT: JACKMAN, JANET
; APPLICANT: SCHOENFELD, JILL R.
; APPLICANT: WILLIAMS, P. MICKEY
; APPLICANT: WOOD, WILLIAM I.
; APPLICANT: WU, THOMAS D.
; TITLE OF INVENTION: Compositions and Methods for the Treatment of Immune
; TITLE OF INVENTION: Related Diseases
; FILE REFERENCE: P1948R1-US
; CURRENT APPLICATION NUMBER: US/10/370,715B
; CURRENT FILING DATE: 2003-02-21
; NUMBER OF SEQ ID NOS: 742
; SEQ ID NO 590
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Homo sapien
; US-10-370-715B-590

Query Match      29.3%; Score 449; DB 5; Length 302;
Best Local Similarity 36.2%; Pred. No. 3 6e-36;
Matches 96; Conservative 52; Mismatches 99; Indels 18; Gaps 3;

QY 36 BEIANFVRPSDVMTVTPKSGTSLQEVVYLVSGADPDEIGLMNIDEQLPVLEYPOP- 94
Db 36 DKINWFQAKPDDLIIISTYPKAGTTWTQEIIVELIQNEGVDEKSKRAPHQRPFFLEMKIPS 95
QY 95 ---GLDIIKELTSPLIKSHLPYRPLPSDLHNGDSKVLYMARNPKDLVSYVYQFHSRLT 151
Db 96 LGSGLQEAHAMPSPRIILKTHLPFLLPSPLEKNCKIIYVARNPKNMVSYYHFORMNKA 155
QY 152 MSYRGTFQFCRRFMNDKLGYSWFHVEHWMDSNVLFKYEDMHRDLVTMVEOLA 211
Db 156 LPAPGTWEEYFETFLAGVCWGSWHEHVKGWEAKDKHRLVLYFYEDMKKPKHIEIQKLA 215
QY 212 RFLGVSCDKAQLEALTEHCHQLV---DQCCNAEALP-----VGRGRVGLWKDIF 257
Db 216 EFIGKLLDKVLDKIVHYTSFDMKQNPMPANYSSIPAEIMDHSISPPFMRKGAVGDKKHIF 275
QY 258 TVSMNEKFDLVYKQMGKCDLTFDF 282
Db 276 TVAQNERFDEYKXKMTDTRLTFHF 300

RESULT 21
US-10-072-012-668
; Sequence 668, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
```

APPLICANT: Padigaru, Muralidhara
APPLICANT: Anderson, David W.
APPLICANT: Rastelli, Luca
APPLICANT: Miller, Charles E.
APPLICANT: Gerlach, Valerie
APPLICANT: Taupier Jr, Raymond J.
APPLICANT: Gusev, Vladimir Y.
APPLICANT: Colman, Steven D.
APPLICANT: Wolenc, Adam R.
APPLICANT: Pena, Carol E. A
APPLICANT: Furtak, Katarzyna
APPLICANT: Grosse, William M.
APPLICANT: Alsobrook II, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-258
CURRENT APPLICATION NUMBER: US/10/072,012
CURRENT FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: 60/265,102
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/265,514
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,517
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,412
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,395
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/266,406
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/266,767
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: 60/267,057
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/266,975
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/267,459
PRIOR FILING DATE: 2001-02-08
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1391
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 668
LENGTH: 296
TYPE: PRT
ORGANISM: Rattus norvegicus
US-10-072-012-668

Query Match 29.2%; Score 447; DB 4; Length 296;
Best Local Similarity 34.4%; Pred. No. 5.6e-36; Mismatches 108; Indels 22; Gaps 5;
Matches 98; Conservative 57;
QY 16 ESKYFEFHVRIPFCRGKMEIANPPVRPSDVWIVTPKSGTSLLOEVVYLVSGADPD 75
DB 10 QTKLEVAGIPLRDSTVDNWSIQIFKAKPDDLICTYPKSGTWTIQEIVNMEQGDVE 69
QY 76 EIGLWNIDEQLVLEY---PQP-GLDIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYM 131
DB 70 KCORTIIQHRHPFIEWARPPQSGVDKANAMPAPRILRTHLPQLLPPSFWTNNCKYLV 129
QY 132 ARNPDLVSYVYQFHSRLTMSYRGTFQFCFRFNDKLGVSFHFVQEFWEHRMDSNV 191
DB 130 ARNAKDCWVSFHYHFRMCVLPNPGTWNBYFTTFNGKRVSGCDFHVKGWWEIIRDRYQI 189
QY 192 LFLKYEDMRDLVTWVEQLARELVGS-----CDKAQLEALTEHCHQLVDQCCNABALP-- 244
DB 190 LFLFYEDMKRDKPREIQKVMQMGKNDLDEVVVDKIVLETSPFEKMKD--NPLTFNSTIPTKT 247
QY 245 -----VGRGRVLWKDITVSNKFDLVYKQKMGKCDLTF 280
DB 248 IMDQISPFMRKIGVGDWKNHFTVAQNERFDEIYEQKMDGTSLNF 292

RESULT 22

US-10-072-012-855
Sequence 855, Application US/10072012
Publication No. US20040033493A1
GENERAL INFORMATION:
APPLICANT: Tchernev, Velizar
APPLICANT: Spytek, Kimberly
APPLICANT: Zerkhusen, Bryan
APPLICANT: Patturajan, Meera
APPLICANT: Shinkets, Richard
APPLICANT: Li, Li
APPLICANT: Gangolli, Esha
APPLICANT: Padigaru, Muralidhara
APPLICANT: Anderson, David W.
APPLICANT: Rastelli, Luca
APPLICANT: Miller, Charles E.
APPLICANT: Gerlach, Valerie
APPLICANT: Taupier Jr, Raymond J.
APPLICANT: Gusev, Vladimir Y.
APPLICANT: Colman, Steven D.
APPLICANT: Wolenc, Adam R.
APPLICANT: Pena, Carol E. A
APPLICANT: Furtak, Katarzyna
APPLICANT: Grosse, William M.
APPLICANT: Alsobrook II, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-258
CURRENT APPLICATION NUMBER: US/10/072,012
CURRENT FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: 60/265,102
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/265,514
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,517
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,412
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,395
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/266,406
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/266,767
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: 60/267,057
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/266,975
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/267,459
PRIOR FILING DATE: 2001-02-08
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1391
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 855
LENGTH: 269
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:
OTHER INFORMATION: Sulfotransferase protein Consensus Sequence
US-10-072-012-855

Query Match 29.0%; Score 444; DB 4; Length 269;
Best Local Similarity 40.6%; Pred. No. 9.9e-36;
Matches 104; Conservative 39; Mismatches 83; Indels 30; Gaps 7;
QY 41 FVVRPSDVWIVTPKSGTSLLOEVVYLVSGAD-----PDEIGLWNIDEQLPVLEYPQPG 95
DB 20 FOARPDVLIAGYPKSGTWTIQEILSLHPNVGDFEPSDDLFRN-----PWLSEYPK-G 73


```

Db      196 DMKNPKKEIKKIRFLKLEKNLNDLRIIHTSFVVMKDNPLVNYTHLPTTVMDSKSP 255
Qy      245 -VGRGRVGLWKDIFTVSNNEKFDLVYKQMKGCDLTF 280
      : : | | | | | | | | | | | | | | | | | | | |
Db      256 PMRGKTAGDKNKNFTVAQNEKFDIAIYETEMSKTALQF 292

RESULT 25
US-10-411-976-13
; Sequence 13, Application US/10411976
; Publication No. US20030170849A1
; GENERAL INFORMATION:
; APPLICANT: Moore, Kevin L.
; TITLE OF INVENTION: TYROSYLPROTEIN SULFOTRANSFERASES
; FILE REFERENCE: 5864.027
; CURRENT APPLICATION NUMBER: US/10/411,976
; PRIOR FILING DATE: 2003-04-11
; PRIOR APPLICATION NUMBER: 09/785,343
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 13
; LENGTH: 295
; TYPE: PRT
; ORGANISM: estrogen sulfotransferase
US-10-411-976-13

Query Match      28.1%; Score 430; DB 4; Length 295;
Best Local Similarity 33.0%; Pred. No. 2.8e-34;
Matches 94; Conservative 57; Mismatches 116; Indels 18; Gaps 3;

Qy      14 EFESKYFEFHGVRLPPFCRCGMKEBIANFPVRPSDVMTVTPKSGTSLLOEVVYLVSQAD 73
      : : | | | | | | | | | | | | | | | | | | | |
Db      7 EYEVFGFGRGLMDKRFYKWEDEVEMFLARPDDLVIATPKSGTTWISVVYIYKEGD 66

Qy      74 PDEIGLMNIDQLPVLEYQPQ-----GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVI 129
      : : | | | | | | | | | | | | | | | | | | | |
Db      67 VEKCKEDAIFNRIPYLECRNEDLINGIKQLKEKESPRIVKTHLPKVLPAFWEKNCCKMI 126

Qy      130 YMAENPKDLVSYVYQFHRSLRTMSYRGTFQEFRCFRFMDKLGYSWFEHVOEFWEHRMDS 189
      : : | | | | | | | | | | | | | | | | | | | |
Db      127 YLCNNAKDVAVSYVYVFLMLMITSYPNPKSFSEFVEKFMQGVPGYSWYDHWKAWKEKSKNS 186

Qy      190 NVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAEALTEHCH-----QLVDQ 236
      : : | | | | | | | | | | | | | | | | | | | |
Db      187 RVLFMFYEDMKEDIRREVVKLIEFLERKPSAELVDRIIQTSTFOEMKNPNSTNYTMPEE 246

Qy      237 CCNAEALP-VGRGRVGLWKDIFTVSNNEKFDLVYKQMKGCDLTF 280
      : : | | | | | | | | | | | | | | | | | | | |
Db      247 MNQKVSFPMRKGIIGDNKNHFPPEALRERFDEHYKQMKDCTVKF 291

RESULT 26
US-10-201-525-13
; Sequence 13, Application US/10201525
; Publication No. US20060009631A1
; GENERAL INFORMATION:
; APPLICANT: Board of Regents of the University of Oklahoma
; TITLE OF INVENTION: TYROSYLPROTEIN SULFOTRANSFERASES AND METHODS OF USE THEREOF
; FILE REFERENCE: 5827.005
; CURRENT APPLICATION NUMBER: US/10/201,525
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/785,343
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: PCT/US99/16750
; PRIOR FILING DATE: 1999-07-23
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 13
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-201-525-13

Query Match      28.1%; Score 430; DB 4; Length 295;
Best Local Similarity 33.0%; Pred. No. 2.8e-34;
Matches 94; Conservative 57; Mismatches 116; Indels 18; Gaps 3;

Qy      14 EFESKYFEFHGVRLPPFCRCGMKEBIANFPVRPSDVMTVTPKSGTSLLOEVVYLVSQAD 73
      : : | | | | | | | | | | | | | | | | | | | |
Db      7 EYEVFGFGRGLMDKRFYKWEDEVEMFLARPDDLVIATPKSGTTWISVVYIYKEGD 66

Qy      74 PDEIGLMNIDQLPVLEYQPQ-----GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVI 129
      : : | | | | | | | | | | | | | | | | | | | |
Db      67 VEKCKEDAIFNRIPYLECRNEDLINGIKQLKEKESPRIVKTHLPKVLPAFWEKNCCKMI 126

Qy      130 YMAENPKDLVSYVYQFHRSLRTMSYRGTFQEFRCFRFMDKLGYSWFEHVOEFWEHRMDS 189
      : : | | | | | | | | | | | | | | | | | | | |
Db      127 YLCNNAKDVAVSYVYVFLMLMITSYPNPKSFSEFVEKFMQGVPGYSWYDHWKAWKEKSKNS 186

Qy      190 NVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAEALTEHCH-----QLVDQ 236
      : : | | | | | | | | | | | | | | | | | | | |
Db      187 RVLFMFYEDMKEDIRREVVKLIEFLERKPSAELVDRIIQTSTFOEMKNPNSTNYTMPEE 246

Qy      237 CCNAEALP-VGRGRVGLWKDIFTVSNNEKFDLVYKQMKGCDLTF 280
      : : | | | | | | | | | | | | | | | | | | | |
Db      247 MNQKVSFPMRKGIIGDNKNHFPPEALRERFDEHYKQMKDCTVKF 291

RESULT 27
US-10-072-012-585
; Sequence 585, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zehrusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangoli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057

```

```
; PRIORITY FILING DATE: 2001-02-07
; PRIORITY APPLICATION NUMBER: 60/266,975
; PRIORITY FILING DATE: 2001-02-07
; PRIORITY APPLICATION NUMBER: 60/267,459
; PRIORITY FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 585
; LENGTH: 307
; TYPE: PRT
; ORGANISM: Gallus gallus
US-10-072-012-585

Query Match      28.0%; Score 428; DB 4; Length 307;
Best Local Similarity 32.6%; Pred. No. 4.8e-34;
Matches 87; Conservative 61; Mismatches 97; Indels 22; Gaps 2;

Qy 36 EEIANFPVRPSDVWITVPKSGTSLLOEVVYLVSGQADPDEIGLNMNIDEQLPVLVYPOP- 94
Db 37 DQVWNFKARPDDLAVATYAKAGTTWTOEIVDMIQNGDIEKCRRASTYKRHPFLEWYIPD 96
Qy 95 -----GLDIIKELTSPRLIKSHLPYRFLPSDLHNGSKVIYMARPNKDLVSVYQFHR 147
Db 97 SSPGLYSGLKAAEAMPSPRTMKTHLPVLQVPPSPFWEQNKIIYVARNAKNDLVSYIHFHR 156
Qy 148 SLRTMSYRGTFQFCRRFMNDKLGYSWFEHVQEFWEHRMDSNVILFLKYEDMHRDLVTMV 207
Db 157 MNKVLDPDGTIEETFEKFWNGEVLGWSYDVKWGWAKDKHRIILYLFEDMKENPKREI 216
Qy 208 EQLARFLGVSCDKAQL-----EALTEHCHQLVDQCCNAEALPVGRGVGLW 253
Db 217 QKIMFLEKDLDEEVLNKKIYNTSPFIMKNDPNMTNYTKDFVGVMDHSVPFMRKGSVGDW 276
Qy 254 KDIFTVSMNEKFDLYVKQKMGKCDLTF 280
Db 277 KNYFTVALNKKPQDQYKKKMDATSLVF 303

RESULT 28
US-10-199-330-5
; Sequence 5, Application US/10199330
; Publication No. US20020182681A1
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-1
; CURRENT APPLICATION NUMBER: US/10/199,330
; PRIOR FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-334-5

Query Match      27.6%; Score 422; DB 4; Length 304;
Best Local Similarity 32.3%; Pred. No. 1.9e-33;
Matches 97; Conservative 59; Mismatches 124; Indels 20; Gaps 5;

Qy 1 MAESEAETPSTGGEFE-SKYPEFHGVRLPPFCRGMGEIAPFVRPSDVWITVPKSGTS 59
Db 1 MAKIEKNAPTWEKPELNFNMEVDGVPPTLIILSKWEKVCNFOAKPDDLIATYPKSGTT 60
Qy 60 LLOEVVYLVSGQADPDEIGLNMNIDEQLPVLV-YP---OPGLDIIKELTSPRLIKSHLPY 114
Db 61 WMHEILDMLNDGVEKCKRAQTLDRHAFLELKPHEKPDLEFVLEWSSQLIKTHLPS 120
Qy 115 RFLPSDLHNGSKVIYMARPNKDLVSVYQFHRSLRTMSYRGTFQFCRRFMNDKLGYS 174
Db 121 HLIPIPSIWKENCKIYVARNPKDCLVSYIHFHRMASFMPDPQNLLEEFYEFKFMGKVVGS 180
Qy 175 WFEHVQEFWEHRMDSNVILFLKYEDMHRDLVTMVQELARFLGVSCDKAQLTEHCHQLV 234
Db 181 WFDHVGMWAAKDMHRILYLFYEDIKKDKREIKLFLKLEKDISEILNKIIYHTSFDV 240
Qy 235 ---DQCCNAEALPVG-----RGRVGLWKDIFTVSMNEKFDLYVKQKMGKCDLTF 280
Db 241 MKQNPMTNYTTLPTSIMDHSISPFMRKGMGPDWKNYFTVAQNEEFDKDYQKMGAGSTLTF 300

RESULT 29
US-10-199-334-5
; Sequence 5, Application US/10199334
; Publication No. US20030008354A1
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-3
; CURRENT APPLICATION NUMBER: US/10/199,334
; PRIOR FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-334-5

Query Match      27.6%; Score 422; DB 4; Length 304;
Best Local Similarity 32.3%; Pred. No. 1.9e-33;
Matches 97; Conservative 59; Mismatches 124; Indels 20; Gaps 5;

Qy 1 MAESEAETPSTGGEFE-SKYPEFHGVRLPPFCRGMGEIAPFVRPSDVWITVPKSGTS 59
Db 1 MAKIEKNAPTWEKPELNFNMEVDGVPPTLIILSKWEKVCNFOAKPDDLIATYPKSGTT 60
Qy 60 LLOEVVYLVSGQADPDEIGLNMNIDEQLPVLV-YP---OPGLDIIKELTSPRLIKSHLPY 114
Db 61 WMHEILDMLNDGVEKCKRAQTLDRHAFLELKPHEKPDLEFVLEWSSQLIKTHLPS 120
Qy 115 RFLPSDLHNGSKVIYMARPNKDLVSVYQFHRSLRTMSYRGTFQFCRRFMNDKLGYS 174
Db 121 HLIPIPSIWKENCKIYVARNPKDCLVSYIHFHRMASFMPDPQNLLEEFYEFKFMGKVVGS 180
Qy 175 WFEHVQEFWEHRMDSNVILFLKYEDMHRDLVTMVQELARFLGVSCDKAQLTEHCHQLV 234
Db 181 WFDHVGMWAAKDMHRILYLFYEDIKKDKREIKLFLKLEKDISEILNKIIYHTSFDV 240
Qy 235 ---DQCCNAEALPVG-----RGRVGLWKDIFTVSMNEKFDLYVKQKMGKCDLTF 280
Db 241 MKQNPMTNYTTLPTSIMDHSISPFMRKGMGPDWKNYFTVAQNEEFDKDYQKMGAGSTLTF 300

RESULT 30
```

```
US-10-199-329-5
; Sequence 5, Application US/10199329
; Publication No. US20030166189A1
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-2
; CURRENT APPLICATION NUMBER: US/10/199,329
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-329-5

Query Match      27.6%; Score 422; DB 4; Length 304;
Best Local Similarity 32.3%; Pred. No. 1.9e-33;
Matches 97; Conservative 59; Mismatches 124; Indels 20; Gaps 5;

Qy 1 MAESEAETPTSGEPE-SKYFEFHGVRLPPFCRCGMEEIANFPVRPSPDVMTVTPKSGTS 59
Db 1 MAKIEKNAPTWEKKPELFNIMEVDGVPPTLILSKWEKVCNFKQAKPDDLLIATPKSGTT 60
Qy 60 LLOEVVYLVSGADPDDEIGLNMNIDEQLPVLE--YP---QPGLDIIKELTSPRLIKSHLPY 114
Db 61 WMHEILMDLNDGDVEKCKRAQTLDRHAFLELKPHEKPDLEFVLEMSSPQLIKTHLPS 120
Qy 115 RFLPSDLHNGDSKVIYMARNPDLVSYQFHRSLRTMSYRGTFQBFRCRRFMNDKLGVS 174
Db 121 HLPPSIWKENCKIYVARNPKDCLVSYHFRMASFMPDPQNLEEFYEKFMGKVVVGS 180
Qy 175 WFEHVQEFWEHRMDSNVLFKLYEDMDHRLDVTMVEQLARFLGVCDCAKQALEALTEHCHLV 234
Db 181 WFDHVGWAAKMDHRIYLFYEDIKDKPKREIEKILKFLKEDISEILNKIIYHTSPDV 240
Qy 235 ---DQCCNAEALPVG-----RGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 241 MKQNPMTNTYTLPTSIMDHSISPFMRKMGPDGKNYFTVAQNEEFDKDYQKKMAGSTLTF 300

RESULT 31
US-11-108-875-5
; Sequence 5, Application US/11108875
; Publication No. US20050196802A1
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-4
; CURRENT APPLICATION NUMBER: US/11/108,875
; CURRENT FILING DATE: 2005-04-19
; PRIOR APPLICATION NUMBER: 10/199,334
; PRIOR FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; NUMBER OF SEQ ID NOS: 23

US-10-199-329-5
; Sequence 5, Application US/10199329
; Publication No. US20030166189A1
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-2
; CURRENT APPLICATION NUMBER: US/10/199,329
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-329-5

Query Match      27.6%; Score 422; DB 4; Length 304;
Best Local Similarity 32.3%; Pred. No. 1.9e-33;
Matches 97; Conservative 59; Mismatches 124; Indels 20; Gaps 5;

Qy 1 MAESEAETPTSGEPE-SKYFEFHGVRLPPFCRCGMEEIANFPVRPSPDVMTVTPKSGTS 59
Db 1 MAKIEKNAPTWEKKPELFNIMEVDGVPPTLILSKWEKVCNFKQAKPDDLLIATPKSGTT 60
Qy 60 LLOEVVYLVSGADPDDEIGLNMNIDEQLPVLE--YP---QPGLDIIKELTSPRLIKSHLPY 114
Db 61 WMHEILMDLNDGDVEKCKRAQTLDRHAFLELKPHEKPDLEFVLEMSSPQLIKTHLPS 120
Qy 115 RFLPSDLHNGDSKVIYMARNPDLVSYQFHRSLRTMSYRGTFQBFRCRRFMNDKLGVS 174
Db 121 HLPPSIWKENCKIYVARNPKDCLVSYHFRMASFMPDPQNLEEFYEKFMGKVVVGS 180
Qy 175 WFEHVQEFWEHRMDSNVLFKLYEDMDHRLDVTMVEQLARFLGVCDCAKQALEALTEHCHLV 234
Db 181 WFDHVGWAAKMDHRIYLFYEDIKDKPKREIEKILKFLKEDISEILNKIIYHTSPDV 240
Qy 235 ---DQCCNAEALPVG-----RGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 241 MKQNPMTNTYTLPTSIMDHSISPFMRKMGPDGKNYFTVAQNEEFDKDYQKKMAGSTLTF 300

RESULT 32
US-09-898-570-40
; Sequence 40, Application US/09898570
; Patent No. US20020123612A1
; GENERAL INFORMATION:
; APPLICANT: GERLACH, VALERIE L.
; APPLICANT: ELLERMAN, KAREN
; APPLICANT: MACDOUGALL, JOHN R.
; APPLICANT: SMITHSON, GLENDA
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND
; TITLE OF INVENTION: METHODS OF USING THE SAME
; FILE REFERENCE: 15966-776CIP
; CURRENT APPLICATION NUMBER: US/09/898,570
; CURRENT FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 60/198,293
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: 60/198,645
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: 60/210,809
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/199,476
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/200,025
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/224,610
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: 60/200,024
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/199,880
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/218,591
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 60/271,814
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/215,855
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 09/839,446
; PRIOR FILING DATE: 2001-04-19
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
```


[illegible][illegible]

```
Qy 81 NIDEQLPVLEY--POP---GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNP 135
Db 61 NTYDRHPFIEWTLPSPLNSGLDLANKMPSPTLTKLHPVHMLPPSPFWKENSIIYVARNA 120
Qy 136 KDLVSYVYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVOEFWEHRMDSNVFLK 195
Db 121 KDLVSYVYFMRNMKMLPDGTLGTYIEQFKAGKVLGWSYDHWKGMWDVQDQHRILYLF 180
Qy 196 YEDMHRDLVTMVEQLARFLGVSCDKAQLAEALTEHCHQLV---DOCCNAEALP----- 244
Db 181 YEDMKEDPKREIKKIAKFLKDISSEVLNKIIYHVSFDMKENPMANVTTLTLPSSIMDHSI 240
Qy 245 ---VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 241 SPFMRKMGPDWKNYFTVAQSEDFEDYRRKMAGSNITF 279

RESULT 38
US-10-199-329-8
; Sequence 8, Application US/10199329
; Publication No. US20030166189A1
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; FILE REFERENCE: CLO00669DIV-2
; CURRENT APPLICATION NUMBER: US/10/199,329
; PRIOR FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-329-8
Query Match 27.4%; Score 419.5; DB 4; Length 283;
Best Local Similarity 32.6%; Pred. No. 3e-33;
Matches 91; Conservative 57; Mismatches 112; Indels 19; Gaps 4;
Qy 21 EFHGVRLPPFCRGKMEETIANFPVRPSDVWIVTYPKSGTSLLOEVVYLVSQGADPDDEIGLM 80
Db 1 EVNGILMSKLSMDNDWKIWNFOAKPDDLLIATYAKAGTTWTQEIYVDMIQNDGDVQKQORA 60
Qy 81 NIDEQLPVLEY--POP---GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNP 135
Db 61 NTYDRHPFIEWTLPSPLNSGLDLANKMPSPTLTKLHPVHMLPPSPFWKENSIIYVARNA 120
Qy 136 KDLVSYVYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVOEFWEHRMDSNVFLK 195
Db 121 KDLVSYVYFMRNMKMLPDGTLGTYIEQFKAGKVLGWSYDHWKGMWDVQDQHRILYLF 180
Qy 196 YEDMHRDLVTMVEQLARFLGVSCDKAQLAEALTEHCHQLV---DOCCNAEALP----- 244
Db 181 YEDMKEDPKREIKKIAKFLKDISSEVLNKIIYHVSFDMKENPMANVTTLTLPSSIMDHSI 240
Qy 245 ---VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 241 SPFMRKMGPDWKNYFTVAQSEDFEDYRRKMAGSNITF 279

US-10-199-329-8
; Sequence 8, Application US/10199329
; Publication No. US20030166189A1
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; FILE REFERENCE: CLO00669DIV-2
; CURRENT APPLICATION NUMBER: US/10/199,329
; PRIOR FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-329-8
Query Match 27.4%; Score 419.5; DB 4; Length 283;
Best Local Similarity 32.6%; Pred. No. 3e-33;
Matches 91; Conservative 57; Mismatches 112; Indels 19; Gaps 4;
Qy 21 EFHGVRLPPFCRGKMEETIANFPVRPSDVWIVTYPKSGTSLLOEVVYLVSQGADPDDEIGLM 80
Db 1 EVNGILMSKLSMDNDWKIWNFOAKPDDLLIATYAKAGTTWTQEIYVDMIQNDGDVQKQORA 60
Qy 81 NIDEQLPVLEY--POP---GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNP 135
Db 61 NTYDRHPFIEWTLPSPLNSGLDLANKMPSPTLTKLHPVHMLPPSPFWKENSIIYVARNA 120
Qy 136 KDLVSYVYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVOEFWEHRMDSNVFLK 195
Db 121 KDLVSYVYFMRNMKMLPDGTLGTYIEQFKAGKVLGWSYDHWKGMWDVQDQHRILYLF 180
Qy 196 YEDMHRDLVTMVEQLARFLGVSCDKAQLAEALTEHCHQLV---DOCCNAEALP----- 244
Db 181 YEDMKEDPKREIKKIAKFLKDISSEVLNKIIYHVSFDMKENPMANVTTLTLPSSIMDHSI 240
Qy 245 ---VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 241 SPFMRKMGPDWKNYFTVAQSEDFEDYRRKMAGSNITF 279

RESULT 39
US-10-199-329-9
; Sequence 9, Application US/10199329
```

```
; Publication No. US20030166189A1
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; FILE REFERENCE: CLO00669DIV-2
; CURRENT APPLICATION NUMBER: US/10/199,329
; PRIOR FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-329-9
Query Match 27.4%; Score 419.5; DB 4; Length 283;
Best Local Similarity 32.6%; Pred. No. 3e-33;
Matches 91; Conservative 57; Mismatches 112; Indels 19; Gaps 4;
Qy 21 EFHGVRLPPFCRGKMEETIANFPVRPSDVWIVTYPKSGTSLLOEVVYLVSQGADPDDEIGLM 80
Db 1 EVNGILMSKLSMDNDWKIWNFOAKPDDLLIATYAKAGTTWTQEIYVDMIQNDGDVQKQORA 60
Qy 81 NIDEQLPVLEY--POP---GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNP 135
Db 61 NTYDRHPFIEWTLPSPLNSGLDLANKMPSPTLTKLHPVHMLPPSPFWKENSIIYVARNA 120
Qy 136 KDLVSYVYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVOEFWEHRMDSNVFLK 195
Db 121 KDLVSYVYFMRNMKMLPDGTLGTYIEQFKAGKVLGWSYDHWKGMWDVQDQHRILYLF 180
Qy 196 YEDMHRDLVTMVEQLARFLGVSCDKAQLAEALTEHCHQLV---DOCCNAEALP----- 244
Db 181 YEDMKEDPKREIKKIAKFLKDISSEVLNKIIYHVSFDMKENPMANVTTLTLPSSIMDHSI 240
Qy 245 ---VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 241 SPFMRKMGPDWKNYFTVAQSEDFEDYRRKMAGSNITF 279

RESULT 40
US-11-108-875-8
; Sequence 8, Application US/11108875
; Publication No. US20050196802A1
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; FILE REFERENCE: CLO00669DIV-4
; CURRENT APPLICATION NUMBER: US/11/108,875
; PRIOR FILING DATE: 2005-04-19
; PRIOR APPLICATION NUMBER: 10/199,334
; PRIOR FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
```



```
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-329-10

Query Match      26.7%; Score 408.5; DB 4; Length 283;
Best Local Similarity 31.5%; Pred. No. 3.9e-32;
Matches 88; Conservative 57; Mismatches 115; Indels 19; Gaps 3;

QY 21 EFHGVRLPPFCRKGMEETANFVRPSDVWIVTPKSGTSLLOEVVYLVSGQADPDPEIGLM 80
Db 1 EVNGILMSKMSNDKIWNFQAKPDDLLIATYAKAGTTWTQEIIVDMIQNDGVDVQKQRA 60
QY 81 NIDEQLPVLEYPOP-----GLDIIKELTSPRIKSHLPYRELPSDLHNGDSKVIYMARNP 135
Db 61 NTYDRHPFIEWTLPPPLNSGLDLANKMPSPTLTKTHLPVQMLPSPFWKENSQIIYVARNA 120
QY 136 KDLVSVYYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVOEFWEHRMDSNVLPFLK 195
Db 121 KDLVSVYYFSSRMKNMLPDGTLGEYIETFKAGKVLGWSWDYHVGWMDVKDKHRLILYLF 180
QY 196 YEDMRDLVTMVEQLARFLGSCDKAQLEALTEHCHQLV---DOCCNAEALP----- 244
Db 181 YEDMKEDPKREIKKIVKFLKDISSEVLNKKIHTSFDMKQNPANVTTLTLPSSIMDHSI 240
QY 245 ---VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 241 SPFMRKGMFGDKWKNYFTVAQSEDFEDYRKQWAGSTITF 279

RESULT 47
US-10-072-012-584
; Sequence 584, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meeta
; APPLICANT: Shinkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072.012
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 584
; LENGTH: 304
```

```
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-072-012-584

Query Match      26.7%; Score 408.5; DB 4; Length 304;
Best Local Similarity 31.5%; Pred. No. 4.3e-32;
Matches 88; Conservative 57; Mismatches 115; Indels 19; Gaps 3;

Qy   21 EFHGVRLPPFCRGKMBEIANPVPSPDVWIVTPKSGTSLQLQEVVYLVSGGADPDDEIGLM 80
Db   22 EVNGILSKMSENMWDKIWFQAOKPDLLIATYAKAGTTWTQBIVDMIQDGDVQCORA 81

Qy   81 NIDGOLPVLEYPOB-----GLDIKELTSPRLIKSHLPVRFPLPSDLHGDSKVYYARNP 135
Db   82 NTYDRHPFIETLDPPLNSGLNDLANKPMSPTLKTHLPVOMLPPSPFWKENSQIIYYARNA 141

Qy   136 KDLVVSYQFHRSLRMTSVYRGTOEQFCRRFMNDKLGYSFEHFVQBFWEHRMDSNVLFKL 195
Db   142 KCLVSYIYSRWNNKMLPDCTLGLEYITFKAGKVLGSWYDHVGWMDVKDKHRLYLIF 201

Qy   196 YEDMRDLVTWVEQLARFLGVSCDKAQLTEHCQLV---DQCCNABALP----- 244
Db   202 YEDMKDPKREIKKIVRFLEKDISEVLNKIIHTTSFDVMKQNPMANYTTLPSSIMDHSI 261

Qy   245 ---VGRGRVGLWKDIFTVSNNKFDLYVKQKMGKCDLTF 280
Db   262 SPFMKGMPGDWKNYFTVAQSEDDEDYRKMGAGSTITF 300

RESULT 48
US-10-199-330-6
; Sequence 6, Application US/10199330
; Publication No. US20020182681A1
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; FILE REFERENCE: CLU00669DIV-1
; CURRENT APPLICATION NUMBER: US/10/199,330
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR FILING DATE: 2000-06-20
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-330-6

Query Match      26.6%; Score 407; DB 4; Length 304;
Best Local Similarity 31.3%; Pred. No. 6e-32;
Matches 94; Conservative 61; Mismatches 125; Indels 20; Gaps 5;

Qy   1 MAESEATPTPGGEFE-SKYFEFHGVRLLPFPCRGKMBEIANFPVRPSDVWIVTPKSGTS 59
Db   1 MAKIEKNAPTMEKKPELFNTIMEVDGPTLLSKWEWKVCNFQAKPDDLIATYPKSGTT 60

Qy   60 LLQEVVYLVSGGADPDDEIGLMNIDEQLPVLE--YP---QPGLDIKELTSPRIKSHLPY 114
Db   61 WMHEILDMLNDGDVEKCKRAQTLDRAHFLKLPFKHKRPDLDFVLEMSSPOLIKTHLPS 120

Qy   115 RFLPSDLHGDSKVYYARNPKDLVSYQFHRSLRMTSVYRGTOEQFCRRFMNDKLGYS 174
Db   121 HLIPPSIWKENCKIYVYARNPKDCLVSYHHFRMASFMPPDNLEEYEFKFMGSKVVGGS 180

Qy   175 WFEHVQBFWEHRMDSNVLFKYEDMRDLVTWVEQLARFLGVSCDKAQLTEHCQLV 234
Db   181 WFDHVKGWMAAKOTHIRILYLFYEDIKKNPKHETHKVLFELEKTLSGDVINIKVHHTSFDV 240

Qy   235 ---DQCCNABALP-----VGRGRVGLWKDIFTVSNNKFDLYVKQKMGKCDLTF 280
Db   241 MKONPMANHTAVPAHIHFNHISIKFMRKGMFGDWKNHFTVAMNENFDKHYEKKWAGSTLNF 300

RESULT 50
US-10-199-329-6
; Sequence 6, Application US/10199329
; Publication No. US20030166189A1
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; FILE REFERENCE: CLU00669DIV-2
; CURRENT APPLICATION NUMBER: US/10/199,329
```

```
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-329-6

Query Match 26.6%; Score 407; DB 4; Length 304;
Best Local Similarity 31.3%; Pred. No. 6e-32;
Matches 94; Conservative 61; Mismatches 125; Indels 20; Gaps 5;

QY 1 MAESEAETPTSGEFE-SKYFEFHGVRLPPFCRGKMEIEIANFPVRPSDVWIVTYPKSGTS 59
D 1 MAKIEKNAPTMEKKPELFNMEVDGVTLLILSKWEWEKVCNFQAKPDDLLIATYPKSGTT 60
QY 60 LLOEVVYLVSQADPDEIGLNMIDQLPVL--YP---QPGLDIIKELTSPRLIKSHLPY 114
D 61 WMHEILDMLNDGVEKCKRAQTLDRHAFLELKFPHKEKPDLEFVLEMSSPQLIKTHLPS 120
QY 115 RFLPSDLHNGSKVIYMARPNKDLVSYQPHRSRLRTMSYRGTFQFCRRFMDKLGYS 174
D 121 HLIPPSIWKENCKIVYVARNPKDCLVSYHFRMASFMPDQNLLEEFYKFMGKVGVS 180
QY 175 WFEHVQEFWEHRMDSNVLFKYEDMHRDLVTVMQARFLGVSCDKAQLALETEHCHOLV 234
D 181 WFDHVKGWAAKOTHRILYLFYEDIKKNPKGHEIHKVLEFLEKTLSGDVINKIVHHTSPDV 240
QY 235 ---DOCCNAEALP-----VGRGRVGLMKDIFTVSMNEKFDLVYKMGKCDLTF 280
D 241 MKDNPMAHNTAVPAHIFNHSISKFMKGMGDKWNHFTVAMNENFDKHYEKKMAGSTLNF 300

RESULT 51
US-11-108-875-6
; Sequence 6, Application US/11108875
; Publication No. US20050196802A1
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-4
; CURRENT APPLICATION NUMBER: US/11/108,875
; CURRENT FILING DATE: 2005-04-19
; PRIOR APPLICATION NUMBER: 10/199,334
; PRIOR FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-108-875-6

Query Match 26.6%; Score 407; DB 6; Length 304;
Best Local Similarity 31.3%; Pred. No. 6e-32;
Matches 94; Conservative 61; Mismatches 125; Indels 20; Gaps 5;

QY 1 MAESEAETPTSGEFE-SKYFEFHGVRLPPFCRGKMEIEIANFPVRPSDVWIVTYPKSGTS 59
D 1 MAKIEKNAPTMEKKPELFNMEVDGVTLLILSKWEWEKVCNFQAKPDDLLIATYPKSGTT 60
QY 60 LLOEVVYLVSQADPDEIGLNMIDQLPVL--YP---QPGLDIIKELTSPRLIKSHLPY 114
D 61 WMHEILDMLNDGVEKCKRAQTLDRHAFLELKFPHKEKPDLEFVLEMSSPQLIKTHLPS 120
QY 115 RFLPSDLHNGSKVIYMARPNKDLVSYQPHRSRLRTMSYRGTFQFCRRFMDKLGYS 174
D 121 HLIPPSIWKENCKIVYVARNPKDCLVSYHFRMASFMPDQNLLEEFYKFMGKVGVS 180
QY 175 WFEHVQEFWEHRMDSNVLFKYEDMHRDLVTVMQARFLGVSCDKAQLALETEHCHOLV 234
D 181 WFDHVKGWAAKOTHRILYLFYEDIKKNPKGHEIHKVLEFLEKTLSGDVINKIVHHTSPDV 240
QY 235 ---DOCCNAEALP-----VGRGRVGLMKDIFTVSMNEKFDLVYKMGKCDLTF 280
D 241 MKDNPMAHNTAVPAHIFNHSISKFMKGMGDKWNHFTVAMNENFDKHYEKKMAGSTLNF 300

RESULT 51
US-11-108-875-6
; Sequence 6, Application US/11108875
; Publication No. US20050196802A1
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-4
; CURRENT APPLICATION NUMBER: US/11/108,875
; CURRENT FILING DATE: 2005-04-19
; PRIOR APPLICATION NUMBER: 10/199,334
; PRIOR FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-108-875-6

Query Match 26.6%; Score 407; DB 6; Length 304;
Best Local Similarity 31.3%; Pred. No. 6e-32;
Matches 94; Conservative 61; Mismatches 125; Indels 20; Gaps 5;

QY 1 MAESEAETPTSGEFE-SKYFEFHGVRLPPFCRGKMEIEIANFPVRPSDVWIVTYPKSGTS 59
D 1 MAKIEKNAPTMEKKPELFNMEVDGVTLLILSKWEWEKVCNFQAKPDDLLIATYPKSGTT 60
QY 60 LLOEVVYLVSQADPDEIGLNMIDQLPVL--YP---QPGLDIIKELTSPRLIKSHLPY 114
D 61 WMHEILDMLNDGVEKCKRAQTLDRHAFLELKFPHKEKPDLEFVLEMSSPQLIKTHLPS 120
QY 115 RFLPSDLHNGSKVIYMARPNKDLVSYQPHRSRLRTMSYRGTFQFCRRFMDKLGYS 174
D 121 HLIPPSIWKENCKIVYVARNPKDCLVSYHFRMASFMPDQNLLEEFYKFMGKVGVS 180
QY 175 WFEHVQEFWEHRMDSNVLFKYEDMHRDLVTVMQARFLGVSCDKAQLALETEHCHOLV 234
D 181 WFDHVKGWAAKOTHRILYLFYEDIKKNPKGHEIHKVLEFLEKTLSGDVINKIVHHTSPDV 240
QY 235 ---DOCCNAEALP-----VGRGRVGLMKDIFTVSMNEKFDLVYKMGKCDLTF 280
D 241 MKDNPMAHNTAVPAHIFNHSISKFMKGMGDKWNHFTVAMNENFDKHYEKKMAGSTLNF 300

RESULT 51
US-11-108-875-6
; Sequence 6, Application US/11108875
; Publication No. US20050196802A1
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-4
; CURRENT APPLICATION NUMBER: US/11/108,875
; CURRENT FILING DATE: 2005-04-19
; PRIOR APPLICATION NUMBER: 10/199,334
; PRIOR FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-108-875-6

Query Match 26.6%; Score 407; DB 6; Length 304;
Best Local Similarity 31.3%; Pred. No. 6e-32;
Matches 94; Conservative 61; Mismatches 125; Indels 20; Gaps 5;

QY 1 MAESEAETPTSGEFE-SKYFEFHGVRLPPFCRGKMEIEIANFPVRPSDVWIVTYPKSGTS 59
D 1 MAKIEKNAPTMEKKPELFNMEVDGVTLLILSKWEWEKVCNFQAKPDDLLIATYPKSGTT 60
QY 60 LLOEVVYLVSQADPDEIGLNMIDQLPVL--YP---QPGLDIIKELTSPRLIKSHLPY 114
D 61 WMHEILDMLNDGVEKCKRAQTLDRHAFLELKFPHKEKPDLEFVLEMSSPQLIKTHLPS 120
QY 115 RFLPSDLHNGSKVIYMARPNKDLVSYQPHRSRLRTMSYRGTFQFCRRFMDKLGYS 174
D 121 HLIPPSIWKENCKIVYVARNPKDCLVSYHFRMASFMPDQNLLEEFYKFMGKVGVS 180
QY 175 WFEHVQEFWEHRMDSNVLFKYEDMHRDLVTVMQARFLGVSCDKAQLALETEHCHOLV 234
D 181 WFDHVKGWAAKOTHRILYLFYEDIKKNPKGHEIHKVLEFLEKTLSGDVINKIVHHTSPDV 240
QY 235 ---DOCCNAEALP-----VGRGRVGLMKDIFTVSMNEKFDLVYKMGKCDLTF 280
D 241 MKDNPMAHNTAVPAHIFNHSISKFMKGMGDKWNHFTVAMNENFDKHYEKKMAGSTLNF 300

RESULT 51
US-11-108-875-6
; Sequence 6, Application US/11108875
; Publication No. US20050196802A1
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-4
; CURRENT APPLICATION NUMBER: US/11/108,875
; CURRENT FILING DATE: 2005-04-19
; PRIOR APPLICATION NUMBER: 10/199,334
; PRIOR FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 304
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-24810

Query Match 26.5%; Score 405; DB 6; Length 316;
Best Local Similarity 34.4%; Pred. No. 1e-31;
Matches 94; Conservative 52; Mismatches 99; Indels 28; Gaps 8;

QY 36 EIEANFPVRPSDVWIVTYPKSGTSLQVYVLSQ-----GADPD-----EI-GLMN 81
D 45 ESIRSLPYQDDVWVMSYPTGSTWAGQWVLLGHLQDLYVAAEQDLRLSLIELSALFS 104
QY 82 IDEQLPVLEYPPQGLDIIKELTSPRLIKSHLPYRFLPSDLHNGSKVIYMARPNKDLVVS 141
D 105 IDHHTVAAKFGNVTDLVRNLPRFRPARSHLPWLLPEQFETVKPRIVYTARNPKDLVVS 164
QY 142 YYQPHRSRLRTMSYRGTFQFCRRFMDKLGYSWFEHVQEFWEHRMDSNVLFKYEDMHR 201
D 142 YYQPHRSRLRTMSYRGTFQFCRRFMDKLGYSWFEHVQEFWEHRMDSNVLFKYEDMHR 201
```



```

QY 60 LLOEVVVLVSQADPDEIGLMNIDQLPVLE--YP---OPGLDIKELTSPRLIKSHLPY 114
Db 61 WHEILDMLNDGDVEKCKRAQTLDRHAFLELKFPHKEKPDLEFVLEMSSPQLIKTHLPS 120
QY 115 RFLPSDLHGDSKVIYMARNPDLVVSYYQFHRSLRTMSYRGTFQFCRRFNMNDKLGYS 174
Db 121 HLIPSIWKENCKIVVARNPKDCLVSYVYHFRMASFMPDPQNLBEFYEKFMGKGFGS 180
QY 175 WFEHQVQFEWHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLEALTEHCHQLV 234
Db 181 WFDHVKGWAAKDMHRILYLFYEDIKQNPKEIHKVLEFLEKTSWGDVINKIVHHTSPDV 240
QY 235 ---DOCCNAEALPV-----GRGRVGLWKDIFTVSNNEKFDLVYKQRMGKCDLT 279
Db 241 MKDNPMANTAVPAHIFNHSISKFMKKGMPGDKWNHFTVALNENFDKHYEKKMAGSTLN 300
QY 280 F 280
Db 301 F 301

```

RESULT 62

```

US-09-839-446-26
; Sequence 26, Application US/09839446
; Publication No. US20030050232A1
; GENERAL INFORMATION:
; APPLICANT: GERLACH, VALERIE L.
; APPLICANT: ELLERMAN, KAREN
; APPLICANT: MACDOUGALL, JOHN R.
; APPLICANT: SMITHSON, GLENDA
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND
; FILE REFERENCE: 15966-776
; CURRENT APPLICATION NUMBER: US/09/839,446
; CURRENT FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: 60/198,293
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: 60/198,645
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: 60/210,809
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/199,476
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/200,025
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/224,610
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: 60/200,024
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/199,880
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/218,591
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 60/271,814
; PRIOR FILING DATE: 2001-02-27
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 26
; LENGTH: 305
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: POLYX
US-09-839-446-26

```

```

Query Match 25.9%; Score 396.5; DB 3; Length 305;
Best Local Similarity 30.9%; Pred. No. 6.9e-31;
Matches 93; Conservative 60; Mismatches 127; Indels 21; Gaps 5;
QY 1 MAESAETPTSGEPE-SKYEFHCHVRLPPFCRGKWEIEIANPVPSPDVWIVTPKSGTS 59
Db 1 MAKIEKNAPTWEKPELNFIMEVDGVPPTLIILSKWEKWCNCFQAKPDDLIATYPRSGTT 60

```

```

QY 60 LLOEVVVLVSQADPDEIGLMNIDQLPVLE--YP---OPGLDIKELTSPRLIKSHLPY 114
Db 61 WHEILDMLNDGDVEKCKRAQTLDRHAFLELKFPHKEKPDLEFVLEMSSPQLIKTHLPS 120
QY 115 RFLPSDLHGDSKVIYMARNPDLVVSYYQFHRSLRTMSYRGTFQFCRRFNMNDKLGYS 174
Db 121 HLIPSIWKENCKIVVARNPKDCLVSYVYHFRMASFMPDPQNLBEFYEKFMGKGFGS 180
QY 175 WFEHQVQFEWHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLEALTEHCHQLV 234
Db 181 WFDHVKGWAAKDMHRILYLFYEDIKQNPKEIHKVLEFLEKTSWGDVINKIVHHTSPDV 240
QY 235 ---DOCCNAEALPV-----GRGRVGLWKDIFTVSNNEKFDLVYKQRMGKCDLT 279
Db 241 MKDNPMANTAVPAHIFNHSISKFMKKGMPGDKWNHFTVALNENFDKHYEKKMAGSTLN 300
QY 280 F 280
Db 301 F 301

```

RESULT 63

```

US-10-381-898-4
; Sequence 4, Application US/10381898
; Publication No. US20040086887A1
; GENERAL INFORMATION:
; APPLICANT: AZIMZAI, Yalda; BAUGHN, Mariah R.;
; APPLICANT: BOROMSKY, Mark L.; DING, Li;
; APPLICANT: DUGAN, Brendan; ELLIOTT, Vicki S.;
; APPLICANT: GANDHI, Ameena R.; GRIFFIN, Jennifer A.;
; APPLICANT: HAFALIA, April J.A.; ISON, Craig H.;
; APPLICANT: KHAN, Farrah A.; LAL, Preeti G.;
; APPLICANT: LEE, Ernestine A.; LU, Dyung Aina M.;
; APPLICANT: ARVIZU, Chandra S.; POLICKY, Jennifer L.;
; APPLICANT: SANJANWALA, Madhusudan M.;
; APPLICANT: TANG, Y. Tom; TRIBOULEY, Catherine M.;
; APPLICANT: CHAWLA, Narinder K.; WALSH, Roderick T.;
; APPLICANT: WARREN, Bridget; XU, Yuming;
; APPLICANT: YANG, Junning; YAO, Monique; YUE, Henry
; TITLE OF INVENTION: DRUG METABOLIZING ENZYMES
; FILE REFERENCE: PI-0233 USN
; CURRENT APPLICATION NUMBER: US/10/381,898
; CURRENT FILING DATE: 2003-10-17
; PRIOR APPLICATION NUMBER: PCT/US01/30662
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US 60/236,947
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: US 60/238,864
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/242,323
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 60/247,581
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/249,519
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: US 60/252,834
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: US 60/250,567
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 71680316CD1
US-10-381-898-4

```

Query Match 25.2%; Score 385.5; DB 4; Length 297;

; PRIOR APPLICATION NUMBER: 60/216,647
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,267
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/216,880
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,270
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/251,869
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/235,834
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/234,274
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/234,223
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/228,924
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/224,518
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,369
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/224,519
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,964
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/241,809
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/249,299
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/236,327
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/241,785
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/244,617
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/225,268
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,368
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/251,856
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/251,868
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/229,344
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/234,997
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: 60/229,343
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,345
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,287
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,513
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/231,413
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/229,509
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/236,367
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/237,039
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,038
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/236,370
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/236,802
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,037
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,040

; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/240,960
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/239,935
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/239,937
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/241,787
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,474
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/246,532
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/249,216
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,210
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/226,681
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,759
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/225,213
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/227,182
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,214
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/235,836
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/230,438
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/215,135
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/225,266
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/249,218
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,208
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,213
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,212
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,207
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,245
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,244
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,217
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,211
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,215
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,264
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,214
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,297
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/232,400
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/231,242
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,081
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,080
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,414
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,244
; PRIOR FILING DATE: 2000-09-08

```

; PRIOR APPLICATION NUMBER: 60/233,064
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/233,063
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,397
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,399
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,401
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/241,808
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,826
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,786
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,221
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,475
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/231,243
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/233,065
; PRIOR FILING DATE: 2000-09-14

Query Match      25.0%; Score 382; DB 3; Length 76;
Best Local Similarity 95.9%; Pred. No. 2.8e-30; Mismatches 1; Indels 0; Gaps 0;
Matches 71; Conservative 1;

Qy 90 EYPOGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYARNPKDLVVSYYQFHRSL 149
Db 1 KYPOGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYARNPKDLVVSYYQFHRSL 60

Qy 150 RTMSYRGTFQEFQR 163
Db 61 RTMSYRGXFQEFQR 74

RESULT 66
US-10-756-149-5061
; Sequence 5061, Application US/10756149
; Publication No. US20050181375A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Zlotnick, Albert
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
; TITLE OF INVENTION: METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER
; FILE REFERENCE: file
; CURRENT APPLICATION NUMBER: US/10/756,149
; CURRENT FILING DATE: 2004-01-12
; NUMBER OF SEQ ID NOS: 5818
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 5061
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-756-149-5061

Query Match      24.9%; Score 381; DB 5; Length 350;
Best Local Similarity 33.3%; Pred. No. 3e-29;
Matches 94; Conservative 52; Mismatches 108; Indels 28; Gaps 7;

Qy 12 PGFESKYFEHGVRLPPFCRG--KMEIA----NFPVRPSDVMIVTYPKSGTSLLOEVV 65
Db 13 PGE----YFRYKGV---PPFVGLYSLSISLAENTQDVRDDDIPIITYPKSGTWTWIEII 65
Qy 66 YLVSGADPDDEIGLNMNIDEQLPVLEYPOPGLDIIKELTSPRLIKSHLPYRFLPSDLHNGD 125
Db 66 CLIIKEGDPWSIRVPIWERAPEW-TIVGAFSIPDQYSPRLMSHLDIQTIFKAFSSK 124
Qy 126 SKVIYARNPKDLVVSYYQFHRSLRTMSYRGTFQEFRCRRFMNDKLGYSWFHVEH 185
Db 125 AKVIYMGNRPRDVSIVSLYHSKIAGQLKDPGTTPDQFLRDFLKGEVQFGSWFDHIKGLMRM 184
```

```

Qy 186 RMDSNVLFXYEDMHRDLVTMVEQLARFLGVSCDKAOLEALTEH---CHQLVDCCNAEA 242
Db 185 KGDNFLFITYEELQQDLQSGVERICGFLGPLGKEALGSVVAHSTFSAMKANTMSNTYL 244
Qy 243 LP-----VGRGRVGLWKDIFTVSMNEKFDLVYKQKM 273
Db 245 LPPSLLDHRRGAFRLKGVCGDKWKNHFTVAQSEAFDRAYRQKM 286
```

RESULT 67

```

US-11-097-143-11343
; Sequence 11343, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11343
; LENGTH: 338
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-11343
```

```

Query Match      23.6%; Score 360.5; DB 6; Length 338;
Best Local Similarity 31.0%; Pred. No. 3.3e-27;
Matches 89; Conservative 51; Mismatches 112; Indels 35; Gaps 7;
```

```

Qy 9 PSTP--GEFESKYFEHGVRLPPFCRGKMEIEIANFPVRPSDVMIVTYPKSGTSLLOEVV 66
Db 34 PSVPVGVNGWEQRF-----CRLADTFQVLDREVDFEVRDDDDVWIVTLFCKGTTWMLQELAW 88
Qy 67 LVSGADPDDEIGLNMNIDEQLPVLEY-----POPGLDIIKELTSPRLIKSHLPYRFLPS 119
Db 89 LVINECDFETAKSVDLTHRSPPFLFNGVVPVNPVPHDTTAAANALSPRLIKSHLPANMLPR 148
Qy 120 DLHNGDSKVIYARNPKDLVVSYYQFHRSLRTMSYRGTFQEFRCRRFMNDKLGYSWFHVEH 179
Db 149 QIWSKRPKIIVVYRNPKDAALSYFHHWRGM--VGYYQTGKDFMHSFIDGVYVNFVTPCWPFI 206
Qy 180 QEFWEHRMDSNVLPKYEDMHRDLVTMVEQLARFLGVSCDKAOLEALTEHC--HQLVDO-- 236
Db 207 LDFMQLRHEPNIFFTSYERMKGQLGVIVSEVAQFLERSVSEQOQMQORHLSFESMRDNP 266
Qy 237 CCN-----AALPVGRGVGLMKDIFTVSMNEKFDL 267
Db 267 ACHVKEFESMKAAAGREVEEFRRVRGVGVGSHKDELTAIIRFDL 313
```

RESULT 68

```
US-11-097-143-14628
; Sequence 14628, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; PRIOR FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14628
; LENGTH: 346
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-14628

Query Match      23.5%; Score 359.5; DB 6; Length 346;
Best Local Similarity 32.3%; Pred. No. 4.2e-27;
Matches 82; Conservative 35; Mismatches 88; Indels 49; Gaps 5;

QY 22 FHGVR-----LPPFCRGKMEETANFPVRPSDVWIVTPKSGTSLQEVVLYVS 69
Db 27 FGERGTGVQVSGEYFFPHKVKDEARYNFEARPDVWIATVPSGTTWTQELIWLVA 86
QY 70 QGADPDEIGLNMIDQLPVLEYP-----QPGLDIHK 100
Db 87 NGLDPEHAQERPLTERFPFFPFPLFVHPKIKEELOENRDSABALEFIEKIAKPGVEALS 146
QY 101 EL--TSPRLIKSHLPYRFLPSDLHNGDSKVIYMARPNKDLVVSYYQFPHRSIRLMTSYRGTF 158
Db 147 EIPRSQRRIKTHFPFSLMPPSVLEKKCKVIYVRDPKDVAVSYVHLNRLFRTOGVGDF 206
QY 159 QSFCCRRFNMNKLGYGSW---FEHVQEFWEHRMDSNVFLFKYEDMHRDLVTWVEQLARFLG 215
Db 207 ERYWYFQN---GLNPWLPYSHVKEAREHAHLSNVLFURYEDMLADLPAINSTIASFLIE 263
QY 216 VSCDKAQLALETHE 229
Db 264 CPPKPEDMDRLLDH 277

RESULT 69
US-11-152-635-11
; Sequence 11, Application US/11152635
; Publication No. US20060024708A1
; GENERAL INFORMATION:
; APPLICANT: Sinclair, James
; APPLICANT: Sinclair, Philip
; TITLE OF INVENTION: Porcine Sulfotransferase 2A1 Polynucleotide Sequence, Protein, an
; TITLE OF INVENTION: Methods of Use for Same
; FILE REFERENCE: P06815US01
; CURRENT APPLICATION NUMBER: US/11/152,635
; CURRENT FILING DATE: 2005-06-14
```

```
; PRIOR APPLICATION NUMBER: 60/580,540
; PRIOR FILING DATE: 2005-06-17
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-11-152-635-11

Query Match      23.3%; Score 356.5; DB 6; Length 285;
Best Local Similarity 31.0%; Pred. No. 6.5e-27;
Matches 83; Conservative 59; Mismatches 109; Indels 17; Gaps 5;

QY 22 FHGVLPP--FCRGKMEIAN-FPVVRPSDVWIVTPKSGTSLQEVVLYVSGADPDEIG 78
Db 8 FEGIAFPTMGFRSETLRKVRDEFVIRDEVDVILTPKSGTNWLABILCLMHKSGDAKWIQ 67
QY 79 LWNIDEQLPVLEYPOPLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARPNKDL 138
Db 68 SVPIWERSPWVE-SEIGYTALSETSPRLFSHLPIQLFPKSFSSKAKVIYLMRNPRDV 126
QY 139 VVSYYQFPHRSIRLMTSYRGTFQFCRRFMNDKLGYSWFEHVQEFWEHRMDSNVFLFKYED 198
Db 127 LVSGYFFWKNMKFIKKPKSWEEYFENFCQGTVVYGSNFDHGHGMPREEKNFLLLSYEE 186
QY 199 MHRDLVTWVEQLARFLGVSCDKAQLALETHEC---HQLVDQCCNAEALPVG----- 246
Db 187 LKQDTGRTIEKICQLFGKLTLEPELNLILKNSSFSQSMKENKMSYLLSDVYVVDKAQLL 246
QY 247 -RGRVGLWKDIFTVSMNEKFDLVYQKMK 273
Db 247 RKGSGDWKNHFTVAQAEFDKLFQEKM 274

RESULT 70
US-10-851-921-12
; Sequence 12, Application US/10851921
; Publication No. US20040241737A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Khawaja, Xavier
; APPLICANT: Xu, Jun
; APPLICANT: Liang, JinJun
; TITLE OF INVENTION: METHODS FOR DIAGNOSING MOOD DISORDERS
; FILE REFERENCE: AM101227
; CURRENT APPLICATION NUMBER: US/10/851,921
; CURRENT FILING DATE: 2004-05-21
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-851-921-12

Query Match      23.2%; Score 355.5; DB 5; Length 285;
Best Local Similarity 31.0%; Pred. No. 8.2e-27;
Matches 83; Conservative 59; Mismatches 109; Indels 17; Gaps 5;

QY 22 FHGVLPP--FCRGKMEIAN-FPVVRPSDVWIVTPKSGTSLQEVVLYVSGADPDEIG 78
Db 8 FEGIAFPTMGFRSETLRKVRDEFVIRDEVDVILTPKSGTNWLABILCLMHKSGDAKWIQ 67
QY 79 LWNIDEQLPVLEYPOPLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARPNKDL 138
Db 68 SVPIWERSPWVE-SEIGYTALSETSPRLFSHLPIQLFPKSFSSKAKVIYLMRNPRDV 126
QY 139 VVSYYQFPHRSIRLMTSYRGTFQFCRRFMNDKLGYSWFEHVQEFWEHRMDSNVFLFKYED 198
Db 127 LVSGYFFWKNMKFIKKPKSWEEYFENFCQGTVVYGSNFDHGHGMPREEKNFLLLSYEE 186
QY 199 MHRDLVTWVEQLARFLGVSCDKAQLALETHEC---HQLVDQCCNAEALPVG----- 246
```


; PRIOR APPLICATION NUMBER: 60/249,218
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,208
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,213
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,212
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,207
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,245
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,244
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,217
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,211
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,215
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,264
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,214
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,297
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/232,400
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/231,242
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,081
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,080
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,414
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,244
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/233,064
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/233,063
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,397
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,399
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,401
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/241,808
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,826
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,786
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,221
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,475
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/231,243
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/233,065
; PRIOR FILING DATE: 2000-09-14

Query Match 20.3%; Score 310; DB 3; Length 105;
Best Local Similarity 95.1%; Pred. No. 7.5e-23;
Matches 58; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 165 FNDKLGYSWFEHVFQEFWEHRMDSNVLFKYEYEDHMDLVTVVEQLARFLGVSCDKAOLE 224
DB 45 FNDKLGYSWFEHVFQEFWEHRMDSNVLFKYEYEDHMDLVTVVEQLARFLGVSCXIFOLE 104
QY 225 A 225

Db 105 A 105
RESULT 75
US-09-795-926-2
; Sequence 2, Application US/09795926
; Patent No. US2002009486A1
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Waite, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Porter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; FILE REFERENCE: LEX-0144-USA
; CURRENT APPLICATION NUMBER: US/09/795,926
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 303
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-795-926-2
Query Match 19.8%; Score 303.5; DB 3; Length 303;
Best Local Similarity 26.7%; Pred. No. 1.5e-21;
Matches 74; Conservative 61; Mismatches 99; Indels 43; Gaps 9;
QY 19 YFEHGVRLP-PFCRGK-MEEIANPPVRPSPVWIVTPKSGTS-----LLOEVVVLVSQGA 72
DB 27 FTYOGIPYPTMTSETFQALDTFEARHDDIVLASYPKCSNWILHIVSELIYAVSK-- 84
QY 73 DPDEIGLMNIDEQLFPVLEYPQ-PGLDI-----IKELTSPLIKSHLPYRFLPSDLHN 123
DB 85 -----KKYKYPEFPVLECGDSEKQYQRMKGFPSPRILATHLHYDKLPGSIFE 130
QY 124 GDSKVIYMARNPDLVSYQFHRSLRTMVSRYGTFOEFCRPFMDKLGYSWFEHVFQEFW 183
DB 131 NKAKILVIFRNPKDTAVSFLFHNDVPDIPSYGSDWDEFFRQFMKGQVSGRYFDFAINWN 190
QY 184 EHRMDSNVLFKYEYEDHMDLVTVVEQLARFLGVSCDKAOLEALTEHCHQLVDQCCNABAL 243
DB 191 KHLDDNVKFLIYEDLKENLAAGIKQIAEFLGFFUTGTGQIQISV---QSTFQAMRAKSQ 247
QY 244 ----PVG-----RGRVGLWKDIFTVSMNEKFDLVYKQ 271
DB 248 DTHGAVGPFPLFRKGEVGDWKNLFSFIQNEKDEKFE 284
RESULT 76
US-10-364-774-2
; Sequence 2, Application US/10364774
; Publication No. US20030144497A1
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn

```
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grose, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 664
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-072-012-664

Query Match      19.8%; Score 303.5; DB 4; Length 303;
Best Local Similarity 26.7%; Pred. No. 1.5e-21;
Matches 74; Conservative 61; Mismatches 99; Indels 43; Gaps 29

QY   19 YFEFGVGLRP-PFCRGK-MEEIANFPVPDVIIVPKSGTGS-----LIQEVLVYSQA 72
DB   27 FTYGQIGYPITMCTSETFQALDTFEARHDDIIVLASYPKCGSNWIIHVSELIVAVSK-- 84
                                           :|::||::||::||::||::||:
QY   73 DPDEIGLANNIDEQLPVLEYPQ-PCGLDI-----IKELTSPLRIKSHLPYRFLPSDLHN 123
DB   85 -----KKKYPEFPVLEGDSSEKQVRMKGPFSPRLATHLHKLPQSIFE 130
                                           :|::||::||::||::||::||:
QY   124 GDSKIYMARNPXOLVSYVGFHRSLRTMSYRGTFQBFCCRFRMNDKLGYSWFHFVQEFW 183
DB   131 NKAKILVIFRNPKOTASVLFHFNVDVPIISYGSWNDEFFRFQFMKGQVSNGRYDFDAINWN 190
                                           :|::||::||::||::||::||:
QY   184 EHMRDSNVLFKYEDMRDLVTVMVEQLARFLGVSCDKAQALEATETCHQLVDPCCNAEAL 243
DB   191 XHLGDGNVKFILYEDLNKENLAAGIKQTAEFLGGFLTGEQIQTTISV---QSTFOAMRAKSQ 247
                                           :|::||::||::||::||::||:
QY   244 ----PVG-----RGVGLWKDIFTVSMNEKFDELVYKQ 271
DB   248 DTHGAVGPFLPRKGEVGDVNKLFSFIQNQMDEKFKXE 284
                                           :|::||::||::||::||::||:

RESULT 78
US-10-468-125-3
; Sequence 3, Application US/10468125
; Publication No. US20040082061A1
; GENERAL INFORMATION:
; APPLICANT: ASTROMOFF, Anna
```

```

; APPLICANT: AU-YOUNG, Janice
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: DING, Li
; APPLICANT: DUGGAN, Brendan M.
; APPLICANT: FORSYTHE, Ian J.
; APPLICANT: GIETZEN, Kimberly J.
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: LEE, Ernestine A.
; APPLICANT: LU, Yan
; APPLICANT: RICHARDSON, Thomas W.
; APPLICANT: RING, Huijun Z.
; APPLICANT: SANTANWALA, Madhusudan
; APPLICANT: SHARNAKAR, Anita
; APPLICANT: CHAWLA, Narinder K.
; APPLICANT: WARREN, Bridget A.
; APPLICANT: XU, Yuming
; APPLICANT: YUE, Henry
; APPLICANT: ZEBARJADIAN, Yeganeh
; TITLE OF INVENTION: DRUG METABOLIZING ENZYMES
; FILE REFERENCE: PI-0363 USN
; CURRENT APPLICATION NUMBER: US/10/468,125
; CURRENT FILING DATE: 2003-08-15
; PRIOR APPLICATION NUMBER: PCT/US02/04918
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/269,643
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/271,332
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: US 60/276,767
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: US 60/282,077
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: US 60/285,447
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: US 60/287,060
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/288,543
; PRIOR FILING DATE: 2001-05-03
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PERL Program
; SEQ ID NO 3
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 7491172CD1
US-10-468-125-3
```

```

Query Match      19.8%; Score 303.5; DB 4; Length 303;
Best Local Similarity 26.7%; Pred. No. 1.5e-21;
Matches 74; Conservative 61; Mismatches 99; Indels 43; Gaps 9;

QY 19 YEFHGVRLP-PFCRGK-MEEIANPVPSPDVWIVTPKSGTS-----LLQEVVVLVSOQA 72
DB 27 PFTYQGIPIPTMCTSETFQALDTFEARHDDIVLASYPKCGSNWILHIVSELIVAVSK-- 84

QY 73 DPDEIGLNMIDQLPVLVEYPQ-PGLDI-----IKELTSPRLIKSHLPYRFLPSDLHN 123
DB 85 -----KKYKYPEPFLVLECGDSEKYQRMKGFPSPRILATHLHYDKLPGSIFE 130

QY 124 GDSKVIYMARNPKDLVSVYQFHRSLRTMSYRGTFQEFRCRRFMNDKLGYSWFHVBQFW 183
DB 131 NKAKILVFRPKDTAVSFLHFNVDPIPSVGSWDEFFRQFMKGQVSGRYDFDAINWN 190

QY 184 EHRMDSNVLFLKYMHRDLVTMVBQLARFLGVSCDKAQLAEALTEHCHQLVDQCCNAEAL 243
DB 191 KHLGDGNVKFPLYEDKENLAAGIKQIAEFLGFLTGEGIQITISV---QSTFQAMRAKSQ 247

QY 244 ----PVG-----RGRVGLWKDIFTVSMNEKFDLVYKQ 271
DB 248 DTHGAVGPFLFRKGEVDKWNLFSEIQNQEMDEKPKF 284
```

```

RESULT 80
US-10-092-900A-150
; Sequence 150, Application US/10092900A
; Publication No. US20040043382A1
; GENERAL INFORMATION:
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Spytek, Kimberly A.
```

```

RESULT 79
US-11-134-241-2
; Sequence 2, Application US/11134241
; Publication No. US20050287568A1
; GENERAL INFORMATION:
; APPLICANT: Doncho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; FILE REFERENCE: LEX-0144-USA
; CURRENT APPLICATION NUMBER: US/11/134,241
; CURRENT FILING DATE: 2005-05-20
; PRIOR APPLICATION NUMBER: US/10/364,774
; PRIOR FILING DATE: 2003-02-11
; PRIOR APPLICATION NUMBER: US/09/795,926
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 303
; TYPE: PRT
; ORGANISM: homo sapiens
US-11-134-241-2

Query Match      19.8%; Score 303.5; DB 6; Length 303;
Best Local Similarity 26.7%; Pred. No. 1.5e-21;
Matches 74; Conservative 61; Mismatches 99; Indels 43; Gaps 9;

QY 19 YEFHGVRLP-PFCRGK-MEEIANPVPSPDVWIVTPKSGTS-----LLQEVVVLVSOQA 72
DB 27 PFTYQGIPIPTMCTSETFQALDTFEARHDDIVLASYPKCGSNWILHIVSELIVAVSK-- 84

QY 73 DPDEIGLNMIDQLPVLVEYPQ-PGLDI-----IKELTSPRLIKSHLPYRFLPSDLHN 123
DB 85 -----KKYKYPEPFLVLECGDSEKYQRMKGFPSPRILATHLHYDKLPGSIFE 130

QY 124 GDSKVIYMARNPKDLVSVYQFHRSLRTMSYRGTFQEFRCRRFMNDKLGYSWFHVBQFW 183
DB 131 NKAKILVFRPKDTAVSFLHFNVDPIPSVGSWDEFFRQFMKGQVSGRYDFDAINWN 190

QY 184 EHRMDSNVLFLKYMHRDLVTMVBQLARFLGVSCDKAQLAEALTEHCHQLVDQCCNAEAL 243
DB 191 KHLGDGNVKFPLYEDKENLAAGIKQIAEFLGFLTGEGIQITISV---QSTFQAMRAKSQ 247

QY 244 ----PVG-----RGRVGLWKDIFTVSMNEKFDLVYKQ 271
DB 248 DTHGAVGPFLFRKGEVDKWNLFSEIQNQEMDEKPKF 284
```



```
RESULT 82
US-09-795-926-6
; Sequence 6, Application US/09795926
; Patent No. US20020098486A1
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING THE SAME
; FILE REFERENCE: LEX-0144-USA
; CURRENT APPLICATION NUMBER: US/09/795,926
; CURRENT FILING DATE: 2001-02-28
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: US 60/191,849
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 265
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-795-926-6

Query Match 19.1%; Score 292.5; DB 3; Length 265;
Best Local Similarity 27.1%; Pred. No. 1.6e-20;
Matches 70; Conservative 57; Mismatches 90; Indels 41; Gaps 7;

QY 36 BEIANFPVRPSDVMIVTPKSGTS-----LLOEVVVLVSQGADPDEIGLWNIDEQLPVLEY 91
Db 8 QALDTFEARHDDIVLASYPRKGSNNILHIVSELIVAVSK-----KKYKY 51

QY 92 PQ-PGLDI-----IKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPKDLVVS 142
Db 52 PEFPVLECGDSEKYQRMKGFPSPRILATHLHYDKLPGSIFENKAKILVIFRNPKDTAVSF 111

QY 143 YQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVEHQRMSNVLFKLYEDMHRD 202
Db 112 LHFHNDVPDIPSYGSDWDEFFRQFMKGQVSWGRYFDFAINWNKHLDDGNVVKFLYEDLKEN 171

QY 203 LVTWVEQLARFLGVSCKAQLAETHCHQLVDQCCNAEL-----PVG-----RGRVGLW 253
Db 172 LAAGIKQIAEFLGFLTGEQITISV---QSTFQAMRAKSQDTHGAVGPFLLFRKGEVGDW 228

QY 254 KDIFTVMNEKFDLVYKQ 271
Db 229 KNLFSEIQNQEMDEKFE 246

RESULT 83
US-10-364-774-6
; Sequence 6, Application US/10364774
; Publication No. US20030144497A1
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
```

```
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING THE SAME
; FILE REFERENCE: LEX-0144-USA
; CURRENT APPLICATION NUMBER: US/10/364,774
; CURRENT FILING DATE: 2003-02-11
; PRIOR APPLICATION NUMBER: US/09/795,926
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 265
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-364-774-6

Query Match 19.1%; Score 292.5; DB 4; Length 265;
Best Local Similarity 27.1%; Pred. No. 1.6e-20;
Matches 70; Conservative 57; Mismatches 90; Indels 41; Gaps 7;

QY 36 BEIANFPVRPSDVMIVTPKSGTS-----LLOEVVVLVSQGADPDEIGLWNIDEQLPVLEY 91
Db 8 QALDTFEARHDDIVLASYPRKGSNNILHIVSELIVAVSK-----KKYKY 51

QY 92 PQ-PGLDI-----IKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPKDLVVS 142
Db 52 PEFPVLECGDSEKYQRMKGFPSPRILATHLHYDKLPGSIFENKAKILVIFRNPKDTAVSF 111

QY 143 YQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVEHQRMSNVLFKLYEDMHRD 202
Db 112 LHFHNDVPDIPSYGSDWDEFFRQFMKGQVSWGRYFDFAINWNKHLDDGNVVKFLYEDLKEN 171

QY 203 LVTWVEQLARFLGVSCKAQLAETHCHQLVDQCCNAEL-----PVG-----RGRVGLW 253
Db 172 LAAGIKQIAEFLGFLTGEQITISV---QSTFQAMRAKSQDTHGAVGPFLLFRKGEVGDW 228

QY 254 KDIFTVMNEKFDLVYKQ 271
Db 229 KNLFSEIQNQEMDEKFE 246

RESULT 84
US-11-134-241-6
; Sequence 6, Application US/11134241
; Publication No. US20050287568A1
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING THE SAME
; FILE REFERENCE: LEX-0144-USA
```

```
; CURRENT APPLICATION NUMBER: US/11/134,241
; CURRENT FILING DATE: 2005-05-20
; PRIOR APPLICATION NUMBER: US/10/364,774
; PRIOR FILING DATE: 2003-02-11
; PRIOR APPLICATION NUMBER: US/09/795,926
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 265
; TYPE: PRT
; ORGANISM: homo sapiens
US-11-134-241-6

Query Match      19.1%; Score 292.5; DB 6; Length 265;
Best Local Similarity 27.1%; Pred. No. 1.6e-20;
Matches 70; Conservative 57; Mismatches 90; Indels 41; Gaps 7;

QY 36 EBIANFPVRPSDVWVITPKSGTS-----LLQEVVLVSQGADPDDEIGLNMNIDEQLPVLEY 91
DB 8 QALDTEARHDDIVLASYPKCGSNILHIVSELIVAVSK-----KKYKY 51

QY 92 PQ-PCGLDI-----IKELTSPRLIKSHLPYRFLPSPDLHNGDSKVIYMARNPKDLVVSY 142
DB 52 PEPFVLECGDSKYQRMKGFPSPRILATHLHVDKLPGSIFENKAKILVIFRNPKDTAVSF 111

QY 143 YQHRSLRTMSYRGTFQEFRCFRFMDNKLGYSGSWFHEHVOEFWEHRMDSNVFLFLKYEDMHRD 202
DB 112 LHFHNDVPDIPSGVSWDEFFRFQMGQVSGRYFDFAINWNKHLDGDNVVKFLYEDLKEN 171

QY 203 LVTWVEQLARFLGVSCDKAQLEALTEHCHQLVDDQCCNAEAL-----PVG-----RGRVGLW 253
DB 172 LAAGIKQIAEFLGFLTGEQIQTISV---QSTFQAMRAKSQDTHGAVGPFPLFRKGEVGDW 228

QY 254 KDIFTVSMNEKFDLVYKQ 271
DB 229 KNLFSEIQNQMDEKFE 246

RESULT 85
US-10-072-012-665
; Sequence 665, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.

; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 665
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Gallus gallus
US-10-072-012-665

Query Match      18.9%; Score 289.5; DB 4; Length 312;
Best Local Similarity 25.4%; Pred. No. 3.9e-20;
Matches 69; Conservative 68; Mismatches 122; Indels 13; Gaps 6;

QY 20 PEFHGVRUP-PFCRGKM-EETANFPVRPSDVWVITPKSGTSLLOEVVLVSQGADPDEI 77
DB 29 FSYKGVLPVALCSPEVFRAMESFEARSDVDVILAGYPKSGTNWVGQILSDLVATFEKERL 88

QY 78 GLMNI-DEQLPVLEYPQPG---LDIIKELTSPRLIKSHLPYRFLPSPDLHNGDSKVIYMA 132
DB 89 BEKSVNDEELEEFPYLEIGDTEKYEYRMKKLPSRRVILTHLSPEKLPKSFKNKAKILLUI 148

QY 133 RNPDLVVSYQFHRSLRTMSYRGTFQEFRCFRFMDNKLGYSGSWFHEHVOEFWEHRMDSNVYL 192
DB 149 RNPDIATSPFHSNRWSALPSYETWDDFFTAFTKMPWGSYFNYLSEWNKYAADENVM 208

QY 193 FLKYEDMHRDLVTWVEQLARFLGVSCDKAQLEALTEHC--HQLVDQCCNAE----ALPVG 246
DB 209 TITVEELKENQTLGVKNIASPFGISLTGEELRSVIERSSFQSMKENSILKTHGALGSMFLR 268

QY 247 RGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDL 278
DB 269 KGVSDWKNLFPNEBQNEKMDKVFEEIARTKL 300

RESULT 86
US-10-072-012-250
; Sequence 250, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
```


US-11-096-568A-7712

Query Match 17.5%; Score 268.5; DB 6; Length 329;
Best Local Similarity 28.1%; Pred. No. 5.4e-18;
Matches 78; Conservative 45; Mismatches 106; Indels 49; Gaps 9;

QY 41 FVPRSDVMIVTPKSGTSLQEVVVLVSQ----GADPDEIGLMNIDEOLPVLEYP--QP 94
Db 63 FVPRADDIILATQPKCGITLWKALAFIANSRYSYGFSDHPLLTRHPQHVPVPIEIPGAGP 122

QY 95 GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPDKLVVSYQFHRSLRTMSY 154
Db 123 NHTDIHTLPPRLATHMPMSLLPPGTRSVGCGEVVYLCRDPKDALVSRHLFENK----AF 178

QY 155 RGT-----FOEFCRRFMNDKLGYSWFEHVOEFWEHRM--DSNVFLFKYEDMHRDLVT 205
Db 179 QGTNLMDNAFMSFCGFS---PYGPFWDHCLGYWRESIARPDNVFLFKYEIISDVP 234

QY 206 MVEQLARFLGVSCDKAQLEALTEHCHQVLDQCCNAEAL-----PVG----- 246
Db 235 VVRKLAKFLGVPLTE---EESSGVAQEVVVKLCSFETLTSLQVNVQGVVGRHGDKIFMSNS 291

QY 247 ----RGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 292 VFYRKEGVDWANHMSHVWGEELDRIVQKLEGSLVF 329

RESULT 95

US-11-096-568A-7711
; Sequence 7711, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 7711
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(330)
; OTHER INFORMATION: Ceres Seq. ID no. 15174865
US-11-096-568A-7711

Query Match 17.5%; Score 268.5; DB 6; Length 330;
Best Local Similarity 28.1%; Pred. No. 5.4e-18;
Matches 78; Conservative 45; Mismatches 106; Indels 49; Gaps 9;

QY 41 FVPRSDVMIVTPKSGTSLQEVVVLVSQ----GADPDEIGLMNIDEOLPVLEYP--QP 94
Db 64 FVPRADDIILATQPKCGITLWKALAFIANSRYSYGFSDHPLLTRHPQHVPVPIEIPGAGP 123

QY 95 GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPDKLVVSYQFHRSLRTMSY 154
Db 124 NHTDIHTLPPRLATHMPMSLLPPGTRSVGCGEVVYLCRDPKDALVSRHLFENK----AF 179

QY 155 RGT-----FOEFCRRFMNDKLGYSWFEHVOEFWEHRM--DSNVFLFKYEDMHRDLVT 205
Db 180 QGTNLMDNAFMSFCGFS---PYGPFWDHCLGYWRESIARPDNVFLFKYEIISDVP 235

QY 206 MVEQLARFLGVSCDKAQLEALTEHCHQVLDQCCNAEAL-----PVG----- 246
Db 236 VVRKLAKFLGVPLTE---EESSGVAQEVVVKLCSFETLTSLQVNVQGVVGRHGDKIFMSNS 292

QY 247 ----RGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 293 VFYRKEGVDWANHMSHVWGEELDRIVQKLEGSLVF 330

RESULT 96

US-11-096-568A-7710
; Sequence 7710, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 7710
; LENGTH: 340
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(340)
; OTHER INFORMATION: Ceres Seq. ID no. 15174864
US-11-096-568A-7710

Query Match 17.5%; Score 268.5; DB 6; Length 340;
Best Local Similarity 28.1%; Pred. No. 5.7e-18;
Matches 78; Conservative 45; Mismatches 106; Indels 49; Gaps 9;

QY 41 FVPRSDVMIVTPKSGTSLQEVVVLVSQ----GADPDEIGLMNIDEOLPVLEYP--QP 94
Db 74 FVPRADDIILATQPKCGITLWKALAFIANSRYSYGFSDHPLLTRHPQHVPVPIEIPGAGP 133

QY 95 GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPDKLVVSYQFHRSLRTMSY 154
Db 134 NHTDIHTLPPRLATHMPMSLLPPGTRSVGCGEVVYLCRDPKDALVSRHLFENK----AF 189

QY 155 RGT-----FOEFCRRFMNDKLGYSWFEHVOEFWEHRM--DSNVFLFKYEDMHRDLVT 205
Db 190 QGTNLMDNAFMSFCGFS---PYGPFWDHCLGYWRESIARPDNVFLFKYEIISDVP 245

QY 206 MVEQLARFLGVSCDKAQLEALTEHCHQVLDQCCNAEAL-----PVG----- 246
Db 246 VVRKLAKFLGVPLTE---EESSGVAQEVVVKLCSFETLTSLQVNVQGVVGRHGDKIFMSNS 302

QY 247 ----RGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 303 VFYRKEGVDWANHMSHVWGEELDRIVQKLEGSLVF 340

RESULT 97

US-10-767-701-42402
; Sequence 42402, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 42402
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C44289_1.pep
US-10-767-701-42402

Query Match 17.5%; Score 267.5; DB 4; Length 344;
Best Local Similarity 28.4%; Pred. No. 7.2e-18;
Matches 78; Conservative 47; Mismatches 109; Indels 41; Gaps 10;

Db		28	PSEKWLVCQYVQGRWHITQALLOGILTCQKHFEAKOSDII	ILVTPNPKSGTTWLKALVPA	87
Qy		67	LVSQADPEIGLNMIDQPLVLEYPQPGLDIIK-----	ELTSPRIKSHLPYRFLPS	119
Db		88	LINRKFPE-VYSVIIILSCYQSALLVPFLGRSLRSPD	DFQSLSPLRNMTHISHLSPLE	146
Qy		120	DLHNGDSKVIYMARNPOLVSYYPFHRSL---	RTMSYRGITQEFCCRRMNDKLGWSWF	176
Db		147	SVKSSCKILVCCRNPKDMFVSLWHFGKKLAPEETADY--	PIEKAVEAFCCQKFTGGPFW	204
Qy		177	EHVOEFWEHRMD--SNVLFKLYEDMHRDLVTVMVEQLAR	PLGVYS-CDKAQLEALTETHC--H	231
Db		205	DHVLEYWYASLENPNKVLFSVEEPPKKTGTETIKRIAEPL	GCGLVGEEBEVRVAILKLCSPF	264
Qy		232	QLVDOCCNAEA-LPVG-----RGRVLWKDIIFTVSMN	KEKDFLVYKQMGKCDLTF	280
Db		265	SLSLVENVREGKLPSCGMETRAPFRKEVGGMRDITL	TESLAEDVIRTIETSEKFGOSGLKF	322

Search completed: May 17, 2006, 11:20:18
Job time : 190 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 17, 2006, 10:59:32 ; Search time 300 Seconds
(without alignments)
875.681 Million cell updates/sec

Title: US-10-768-158-2

Perfect score: 1530

Sequence: 1 MAESEATPSTPGSFESKYF.....FDLVYKQMGKCDLTFDPYL 284

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : UniProt 7.2.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1530	100.0	284	1	ST4A1_HUMAN
2	1525	99.7	284	2	QSR4X3_PONPY
3	1508	98.6	284	2	Q865L7_RABIT
4	1506	98.4	284	1	ST4A1_MOUSE
5	1506	98.4	284	1	ST4A1_RAT
6	1506	98.4	284	2	Q3TXV5_MOUSE
7	1483	96.9	284	2	Q8BQV8_MOUSE
8	1385	90.5	284	2	Q68EV3_XENLA
9	1382	90.3	284	2	Q52KW3_XENLA
10	1313	85.8	284	2	Q4SRF8_TETNG
11	995	65.0	218	2	Q5RCAB_PONPY
12	495	32.4	298	2	Q9R1S2_MOUSE
13	489	32.0	291	1	ST1A1_MOUSE
14	480	31.4	296	1	ST1C1_MOUSE
15	476	31.1	296	2	Q3UNA9_MOUSE
16	472	30.8	295	1	ST1A3_HUMAN
17	467	30.5	296	1	STK1_RAT
18	467	30.5	296	2	Q3ZAV7_RAT
19	466	30.5	296	2	Q6ZG8_RAT
20	465	30.4	296	2	Q569D0_RAT
21	464	30.3	295	2	Q2TAB3_HUMAN
22	463	30.3	298	2	Q58CV8_BOVIN
23	459.5	30.0	296	1	ST1B1_CHICK
24	458	29.9	291	1	ST1A1_RAT
25	458	29.9	291	2	Q548D2_RAT
26	456	29.8	295	1	ST1A2_HUMAN
27	456	29.8	295	2	Q6Y0X5_PIG
28	456	29.8	296	1	STK2_RAT
29	453.5	29.6	296	2	Q95JC6_TRIVU
30	451	29.5	295	1	ST1A1_HUMAN
31	451	29.5	295	2	Q2NL71_HUMAN

32	451	29.5	295	2	Q3ZC30_BOVIN
33	450	29.4	295	1	ST1A1_CANFA
34	449	29.3	295	2	Q95KM3_RABIT
35	449	29.3	296	1	ST1C1_HUMAN
36	449	29.3	302	1	ST1C2_HUMAN
37	447	29.2	295	2	Q95MF8_PIG
38	446.5	29.2	299	1	ST1B1_RAT
39	446	29.2	295	1	ST1A1_MACFA
40	446	29.2	295	1	ST1E1_BOVIN
41	443.5	29.0	296	1	ST1B1_HUMAN
42	443	29.0	296	1	ST1B1_CANFA
43	442.5	28.9	296	1	ST1C1_RABIT
44	442	28.9	292	2	Q9XT98_RABIT
45	440	28.8	297	2	Q8GP49_XENLA
46	436.5	28.5	296	2	Q3T0Y3_BOVIN
47	435	28.4	295	1	ST1E2_RAT
48	435	28.4	295	1	ST1E6_RAT
49	434.5	28.4	299	1	ST1S1_BRARE
50	434	28.4	294	1	ST1E1_HUMAN
51	434	28.4	294	2	Q53X91_HUMAN
52	432.5	28.3	301	2	Q6DHG7_BRARE
53	432	28.2	295	2	Q9D566_MOUSE
54	431.5	28.2	287	2	Q5HZV7_XENTR
55	431	28.2	295	1	ST1E1_MOUSE
56	431	28.2	295	2	Q8JZX7_MOUSE
57	430	28.1	295	2	Q99ND5_RAT
58	429.5	28.1	299	1	ST1B1_MOUSE
59	429	28.0	263	2	Q91W19_MOUSE
60	429	28.0	295	1	ST1E3_RAT
61	429	28.0	295	2	Q3T0S9_BOVIN
62	428	28.0	307	2	Q90WR6_CHICK
63	427.5	27.9	301	1	ST1S3_BRARE
64	427.5	27.9	301	1	ST2S2_BRARE
65	427.5	27.9	302	2	Q4FZP1_XENLA
66	423	27.6	294	2	Q2XV97_MACFA
67	422	27.6	304	2	Q61M16_HUMAN
68	421.5	27.5	304	2	Q8XZC1_BRARE
69	421	27.5	295	1	ST1E1_RAT
70	419.5	27.4	285	2	Q3HYK0_STRPU
71	419.5	27.4	304	1	SUAC_RAT
72	419.5	27.4	304	2	Q5M8B5_RAT
73	416.5	27.2	287	2	Q6P8G4_XENTR
74	414.5	27.1	276	2	Q68EV4_XENLA
75	414	27.1	295	2	Q95JCS_ORNAN
76	413	27.0	295	2	Q35401_MOUSE
77	413	27.0	303	2	Q5EAW0_XENLA
78	412.5	27.0	304	2	Q80VR3_MOUSE
79	412	26.9	294	1	ST1A1_BOVIN
80	412	26.9	294	2	Q7S293_XENLA
81	412	26.9	295	2	Q95JD6_CANFA
82	412	26.9	295	2	Q3UZ26_MOUSE
83	412	26.9	295	2	Q6NZD1_MOUSE
84	412	26.9	308	2	Q921G0_RAT
85	412	26.9	309	2	Q9R2C2_MOUSE
86	411	26.9	305	2	Q66KW4_XENLA
87	410.5	26.8	307	2	Q86KW4_XENTR
88	408.5	26.7	296	1	ST1E1_CAVPO
89	408.5	26.7	304	2	Q70262_MOUSE
90	405.5	26.5	309	2	Q4RG19_TETNG
91	405	26.5	316	2	Q9VHH0_DROME
92	404.5	26.4	287	2	Q6DJL6_XENLA
93	402	26.3	287	2	Q5M7C2_XENLA
94	401.5	26.2	293	2	Q49IK6_BRARE
95	401	26.2	304	2	Q61M15_HUMAN
96	395.5	25.8	307	2	Q53SG4_HUMAN
97	393.5	25.7	301	1	ST3A1_RABIT
98	393.5	25.7	304	2	Q6DJ68_XENTR
99	393.5	25.7	304	2	Q7ZYH0_XENLA
100	392.5	25.7	299	2	Q6GL93_XENTR

ALIGNMENTS

RESULT 1
ST4A1 HUMAN
ID ST4A1_HUMAN STANDARD; PRT; 284 AA.
AC Q9BR01; O43728;
DC 02-AUG-2002, integrated into UniProtKB/Swiss-Prot.
DT 02-AUG-2002, sequence version 2.
DT 07-FEB-2006, entry version 33.
DE Sulfotransferase 4A1 (EC 2.8.2.-) (Brain sulfotransferase-like
DE protein) (HBR-STL) (HBR-STL-1) (Nervous system sulfotransferase
DE (NST)).
GN Name=SULT4A1; Synonyms=SULTX3;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE (ISOFORM 1), AND CHARACTERIZATION.
RC TISSUE=Brain;
RX MEDLINE=20164074; PubMed=10698717; DOI=10.1042/0264-6021.3460857;
RA Falany C.N., Xie X., Wang J., Ferrer J., Falany J.L.;
RT "Molecular cloning and expression of novel sulphotransferase-like
RT cDNAs from human and rat brain.";
RL Biochem. J. 346:857-864(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE (ISOFORM 1), AND CHARACTERIZATION.
RC TISSUE=Brain;
RX MEDLINE=22035352; PubMed=12039030; DOI=10.1016/S0378-1119(02)00431-6;
RA Sakakibara Y., Suiko M., Pai T.G., Nakayama T., Takami Y.,
RA Katafuchi J., Liu M.-C.;
RT "Highly conserved mouse and human brain sulfotransferases: molecular
RT cloning, expression, and functional characterization.";
RL Gene 285:39-47(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE (ISOFORM 1).
RC TISSUE=Brain;
RA Martin S.C., Farb D.H.;
RT "Molecular identification of a human nervous system cytoplasmic
RT sulfotransferase, NST.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE (ISOFORM 1).
RC TISSUE=Brain;
RA Walther S.E., Rastogi S.B.;
RT "Molecular and physical characterization of human SUL74A1,
RT representing a novel cytosolic sulfotransferase family.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP NUCLEOTIDE SEQUENCE (ISOFORM 2).
RC TISSUE=Testis;
RA Goward M.E., Huckle E.J.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [6]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).
RX PubMed=15461802; DOI=10.1196/gh-2004-5-10-r84;
RA Collins J.E., Wright C.L., Edwards C.A., Davis M.P., Grinham J.A.,
RA Cole C.G., Goward M.E., Aguado B., Mallia M., Mokrab Y., Huckle E.J.,
RA Beare D.M., Dunham I.;
RT "A genome annotation-driven approach to cloning the human ORFeome.";
RL Genome Biol. 5:RESEARCH84.1-RESEARCH84.11(2004).
RN [7]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=20057165; PubMed=10591208; DOI=10.1038/990031;
RA Dunham I., Hunt A.R., Collins J.E., Bruskiewicz R., Beare D.M.,
RA Clamp M., Mink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
RA Bagnoley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,
RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,
RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,
RA Dhani P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,
RA Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,

Gilbert J.G.R., Goward M.E., Grahame D.V., Griffiths M.N.D., Hall C.,
RA Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S.,
RA Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,
RA Laird G.K., Langford C.F., Leversha M.A., Lloyd C., Lloyd D.M.,
RA Martyn I.D., Mashreghi-Mohammadi M., Matthews L.H., McCann O.T.,
RA Mcclay J., McLaren S., McMurray A.M., Milne S.A., Mortimore B.J.,
RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,
RA Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Ross M.T.,
RA Scott C.E., Sehra H.K., Stuce C.D., Smalley S., Smith M.L., Swann R.M.,
RA Soudern L.C., Spraggon L., Steward C.A., Sulston J.E., Swann R.M.,
RA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,
RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,
RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,
RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,
RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyama S.,
RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,
RA Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,
RA Lewis J., Lewis S., Lin S.-P., Loh P., Mala J.E., Nguyen T., Pan H.,
RA Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,
RA Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,
RA Zhan M., Zhang G., Chisoso S., Murray J., Miller N., Minx P.,
RA Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,
RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,
RA Hinds K., Kemp K., Latreille P., Layman D., Ozeresky P., Rohlfing T.,
RA Scheet P., Walker C., Wamaley A., Wohldmann P., Pepin K., Nelson J.,
RA Korf I., Bedell J.A., Hillier L.W., Mardis E., Waterston R.,
RA Wilson R., Emanuel B.S., Shaikh T., Kurahashi H., Saitta S.,
RA Budarf M.L., Mcdermid H.E., Johnson A., Wong A.C.C., Morrow B.E.,
RA Edelmann L., Kim U.J., Shizuya H., Simon M.I., Dumanaki J.P.,
RA Peyrard M., Kedra D., Seroussi E., Francon I., Tapia I., Bruder C.E.,
RA O'Brien K.P., Wilkinson P., Bodenteich A., Hartman K., Hu X.,
RA Khan A.S., Lane L., Tilahun Y., Wright H.;
RT "The DNA sequence of human chromosome 22.";
RL Nature 402:489-495(1999).
RN [8]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORMS 1 AND 2).
RC TISSUE=Fetal brain, Hippocampus, and Hypothalamus;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haiech F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano M.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Rulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903(2002).
CC -!- FUNCTION: May catalyze the sulfate conjugation of many drugs,
CC xenobiotic compounds, hormones, and neurotransmitters. Displays
CC activity towards L-triiodothyronine, thyroxine, estrone, p-
CC nitrophenol, 2-naphthylamine, and 2-naphthol. May have a role in the
CC sulfation of drugs and neurotransmitters in the CNS.
CC -!- SUBCELLULAR LOCATION: Cytoplasm (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q9BR01-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q9BR01-2; Sequence=VSP_006304;
CC Note=No experimental confirmation available;
CC -!- TISSUE SPECIFICITY: Highly expressed in the cerebral cortex and
CC frontal lobe, slightly less in the cerebellum, occipital and

CC temporal lobes, relatively low in the medulla and putamen, and
CC lowest in the spinal cord. No expression detected in the pancreas
CC (Ref.1). Highly expressed in fetal brain and occipital lobe,
CC slightly less in the whole brain, frontal lobe, hippocampus, and
CC lung, very low expression in cerebellum, medulla oblongata,
CC temporal lobe, testis, kidney and appendix (Ref.2).
CC -!- SIMILARITY: Belongs to the sulfotransferase family.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF188698; AAF61197.1; -; mRNA.
DR EMBL; AF115311; AAF21970.1; -; mRNA.
DR EMBL; AF176342; AAK64595.1; -; mRNA.
DR EMBL; AF251263; AAF98152.1; -; mRNA.
DR EMBL; AL590119; CAC34872.1; -; mRNA.
DR EMBL; CR456588; CAG30474.1; -; mRNA.
DR EMBL; Z97055; CAB09788.1; -; Genomic_DNA.
DR EMBL; BC022459; AAH22459.1; -; mRNA.
DR EMBL; BC028171; AAH28171.1; -; mRNA.
DR EMBL; BC030665; AAH30665.1; -; mRNA.
DR HSSP; P50224; 1CJM.
DR Ensembl; ENSG00000130540; Homo sapiens.
DR HGNC; HGNC:14903; SULT4A1.
DR MIM; 608359; gene.
DR GO; GO:0005829; C:cytosol; NAS.
DR GO; GO:0008146; F:sulfotransferase activity; NAS.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase 1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Alternative splicing; Lipid metabolism; Steroid metabolism;
KW Transferase.
FT CHAIN 1 284 Sulfotransferase 4A1.
FT FT FTID=PRO 0000085167.
FT NP_BIND 246 254 PAPS (Potential).
FT VARSPIC 248 284 GRVGLWKDIFTVSMNEKEDLVYKQMGKCDLTFDFYL ->
FT AHCVFARKIFLSW (in isoform 2).
FT FTFTID=VSP 006304.
FT CONFLICT 55 56 KS -> P (in Ref. 8; AAH30665).
FT CONFLICT 239 239 N -> S (in Ref. 8; AAH22459).
SQ SEQUENCE 284 AA; 33085 MW; A6EA6844B66C400B CRC64;

Query Match 100.0%; Score 1530; DB 1; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.2e-119;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAESEAETPTGPEFSKYFEFHGVRLLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60
Db 1 MAESEAETPTGPEFSKYFEFHGVRLLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60

QY 61 LQEVVYLSQADPDEIGLMNIDEQLPVLEYQPQGLDIIKELTSRLIKSHLPYRFLPSD 120
Db 61 LQEVVYLSQADPDEIGLMNIDEQLPVLEYQPQGLDIIKELTSRLIKSHLPYRFLPSD 120

QY 121 LHNGDSKVIYARNPKDLVSVYQFHRSLRTMSYRGTFQFCRRPMNDKLGYSWFHVQ 180
Db 121 LHNGDSKVIYARNPKDLVSVYQFHRSLRTMSYRGTFQFCRRPMNDKLGYSWFHVQ 180

QY 181 EFWEHRMDSNVLFLEYEDMHRDLVTMVQEQLARFLGVSCDKAQLTEHCHQLVDQCCNA 240
Db 181 EFWEHRMDSNVLFLEYEDMHRDLVTMVQEQLARFLGVSCDKAQLTEHCHQLVDQCCNA 240

QY 241 EALPVGGRVGLWKDIFTVSMNEKEDLVYKQMGKCDLTFDFYL 284
Db 241 EALPVGGRVGLWKDIFTVSMNEKEDLVYKQMGKCDLTFDFYL 284

RESULT 2
Q5R4X3_PONPY PRELIMINARY; PRT; 284 AA.
ID Q5R4X3_PONPY
AC Q5R4X3_PONPY
DT 21-DEC-2004, integrated into UniProtKB/TrEMBL.
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License

DT 07-FEB-2006, entry version 7.
DE Hypothetical protein DKFZp459L1114.
GN Name=DKFZp459L1114;
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Cortex;
RG The German cDNA Consortium;
RA Bloeker H., Boecher M., Brandt P., Mewes H.W., Weil B., Amid C.,
RA Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; CR861116; CAH93193.1; -; mRNA.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase 1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Hypothetical protein.
SQ SEQUENCE 284 AA; 33111 MW; 531F6C10A3294B00 CRC64;

Query Match 99.7%; Score 1525; DB 2; Length 284;
Best Local Similarity 99.6%; Pred. No. 3.2e-119;
Matches 283; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAESEAETPTGPEFSKYFEFHGVRLLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60
Db 1 MPSEAEATPTGPEFSKYFEFHGVRLLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60

QY 61 LQEVVYLSQADPDEIGLMNIDEQLPVLEYQPQGLDIIKELTSRLIKSHLPYRFLPSD 120
Db 61 LQEVVYLSQADPDEIGLMNIDEQLPVLEYQPQGLDIIKELTSRLIKSHLPYRFLPSD 120

QY 121 LHNGDSKVIYARNPKDLVSVYQFHRSLRTMSYRGTFQFCRRPMNDKLGYSWFHVQ 180
Db 121 LHNGDSKVIYARNPKDLVSVYQFHRSLRTMSYRGTFQFCRRPMNDKLGYSWFHVQ 180

QY 181 EFWEHRMDSNVLFLEYEDMHRDLVTMVQEQLARFLGVSCDKAQLTEHCHQLVDQCCNA 240
Db 181 EFWEHRMDSNVLFLEYEDMHRDLVTMVQEQLARFLGVSCDKAQLTEHCHQLVDQCCNA 240

QY 241 EALPVGGRVGLWKDIFTVSMNEKEDLVYKQMGKCDLTFDFYL 284
Db 241 EALPVGGRVGLWKDIFTVSMNEKEDLVYKQMGKCDLTFDFYL 284

RESULT 3
Q865L7_RABIT PRELIMINARY; PRT; 284 AA.
ID Q865L7_RABIT
AC Q865L7;
DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2003, sequence version 1.
DE 07-FEB-2006, entry version 11.
DE Brain sulfotransferase 4A1.
GN Name=Sult4a1;
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Lagomorpha; Leporidae;
OC Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA He D., Falany C.N.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License

```

CC -----
DR EMBL; AY196782; AAO45181.1; -, mRNA.
DR HSP; P50224; ICM.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
DR Transferrase.
SQ SEQUENCE 284 AA; 33107 MW; 782B6899407A090DE CRC64;
-----
      Query Match          98.6%; Score 1508; DB 2; Length 284;
      Best Local Similarity 98.2%; Pred. No. 8.6e-118;
      Matches 279; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 MAESEAETPTGPEESKYFPHGVRLPPFCRGKMEETANFVRPSDWMIVTPKSGTSL 60
DB 1 MAESEAETPTGPEESKYFPHGVRLPPFCRGKMEETANFVRPSDWMIVTPKSGTSL 60
QY 61 LOEVVYLYSQADPDPEIGLMNIDEQLPVLEYPOQGLDIKELTSPRLIKSHLPYRFLPSD 120
DB 61 LOEVVYLYSQADPDPEIGLMNIDEQLPVLEYPOQGLDIKELTSPRLIKSHLPYRFLPSD 120
QY 121 LHNGDSKVIYARNPKDLVSVYQPHRSILRTMSYRGTFQFCRRPMNDKLGYSWFHVQ 180
DB 121 LHNGDSKVIYARNPKDLVSVYQPHRSILRTMSYRGTFQFCRRPMNDKLGYSWFHVQ 180
QY 181 EFWEHRMDSNVLFLKYEDHMDRLVTMVEQLARFLGVSCDKAQLEALTECHQOLVDQCCNA 240
DB 181 EFWEHRMDSNVLFLKYEDHMDRLVTMVEQLARFLGVSCDKAQLEALTECHQOLVDQCCNA 240
QY 241 EALPVGRGRVGLWKDIFTVSMNEKEDLVYKQMGKCDLTFEYL 284
DB 241 EALPVGRGRVGLWKDIFTVSMNEKEDLVYKQMGKCDLTFEYL 284

RESULT 4
ST4A1 MOUSE
ID ST4A1_MOUSE STANDARD; PRT; 284 AA.
AC P63046; O98872; Q91X55; Q9CWMY7; Q9DC97;
DT 31-AUG-2004, integrated into UniProtKB/Swiss-Prot.
DT 31-AUG-2004, sequence version 1.
DT 07-FEB-2006, entry version 15.
DE Sulfotransferase 4A1 (EC 2.8.2.-) (Brain sulfotransferase-like
protein) (msr-STL) (Nervous system sulfotransferase) (NST).
GN Name=St4a1; Synonyms=Sultx3;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Murcidae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), FUNCTION, AND TISSUE
RP SPECIFICITY.
RC TISSUE=Brain;
RX MEDLINE=22035352; PubMed=12039030; DOI=10.1016/S0378-1119(02)00431-6;
RA Sakakibara Y., Suiko M., Pai T.G., Nakayama T., Takami Y.,
RA Katafuchi J., Liu M.-C.;
RT "Highly conserved mouse and human brain sulfotransferases: molecular
cloning, expression, and functional characterization.";
RT Gene 285:35-47(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORMS 1 AND 2).
RC STRAIN=C57BL/6J; TISSUE=Embryonic stem cell;
RX PubMed=16141072; DOI=10.1126/science.1112014;
RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,
RA Bajic V.B., Bremner S.E., Batalov S., Forrest A.R., Zavolan M.,
RA Davis M.J., Wilmring L.G., Aidinis V., Allen J.E.,
RA Ambesi-Impombato A., Apweiler R., Aturaliya R.N., Bailey T.L.,
RA Bansal M., Baxter L., Bissel K.W., Bersano T., Bono H., Chalk A.M.,
RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,

```

```

di Bernardo D., Down T., Engstrom P., Fagiolini M., Faulkner G.,
Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,
Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,
Grigucinic S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
Hill D., Huminicki L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,
Kitano H., Kollias G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
Matsuda H., Matsuura S., Miki H., Mignone F., Miyake S., Morris K.,
Mottagui-Tabar S., Mulder N., Nakano N., Nakautchi H., Ng P.,
Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavese G., Pesole G.,
Petrovsky N., Piazza S., Reed J., Reid J.P., Ring B.Z., Ringwald M.,
Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C., Sheng Y.,
Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sheng Y.,
Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
Sperling S., Stupka E., Sugiyara K., Sultana R., Takenaka Y., Taki K.,
Tannoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,
Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,
Wahlestedt S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
Wahlstedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Aragawa T.,
Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Ninomiya N.,
Nishio T., Okada M., Plessey C., Shibata K., Shiraki T., Suzuki S.,
Tagami M., Waki K., Wachihi A., Okamura-Oho Y., Suzuki H., Kawai J.,
Hayashizaki Y.;
RT "The transcriptional landscape of the mammalian genome.";
RN Science 309:1559-1563(2005).
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Roak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: May catalyze the sulfate conjugation of many drugs,
CC xenobiotic compounds, hormones, and neurotransmitters. Displays
CC activity towards 1-trifluoromethyl-2-thiouracil, estrone, p-
CC nitrophenol, 2-naphthylamine, and 2-naphthol. May have a role in the
CC sulfation of drugs and neurotransmitters in the CNS.
CC -!- SUBCELLULAR LOCATION: Cytoplasm (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=P63046-1, Q9DC97-1;
CC Sequence=Displayed;
CC Name=2;
CC IsoId=P63046-2, Q9DC97-2;
CC Sequence=VSP_006305;
CC -!- TISSUE SPECIFICITY: Expressed in brain, cerebellum and
CC hypothalamus. Not detected in pancreas, liver, lung, intestine,
CC kidney, uterus, adrenal gland, thymus, spleen, epididymis,
CC testicle, and heart.

```

```

CC CC -!- SIMILARITY: Belongs to the sulfotransferase family.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF059257; AAC63999.1; -; mRNA.
DR EMBL; AK003034; BAB22522.1; -; mRNA.
DR EMBL; AK010293; BAB26829.1; -; mRNA.
DR EMBL; BC051132; AAH51132.1; -; mRNA.
DR EMBL; BC054757; AAH54757.1; -; mRNA.
DR HSSP; P50224; 1CJM.
DR MGI; MGI:1888971; Sult4a1.
DR GO; GO:0005737; C:cytoplasm; TAS.
DR GO; GO:0008146; F:sulfotransferase activity; IDA.
DR GO; GO:0008790; F:sulfur metabolism; IDA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase 1; 1.
KW Alternative splicing; Lipid metabolism; Steroid metabolism;
KW Transferase.
FT CHAIN 1 284 Sulfotransferase 4A1.
FT NP_BIND 246 254 /FTID=PRO_0000085168.
FT VARSP_LIC 248 284 GRAVLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL ->
FT ACHLFTOKIALRWRCGRSGSLHCLDLVHVTA (in isoform 2).
FT CONFLICT 6 6 /FTID=VSP_006305.
FT SEQUENCE 284 AA; 33054 MW; FCAE940F721956FC CRC64; A -> R (in Ref. 2; BAB22522).
Query Match 98.4%; Score 1506; DB 1; Length 284;
Best Local Similarity 97.9%; Pred. No. 1.3e-117; Indels 0; Gaps 0;
Matches 278; Conservative 4; Mismatches 2;
QY 1 MAESEAETPTGPGFESKYFEFGVRLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60
Db 1 MAESEAETPTGPGFESKYFEFGVRLPPFCRGKMEETADFPVRPSDVWIVTPKSGTSL 60
QY 61 LOEVVYLVSQGADPDEIGLMNIDQLPVLEYVQPGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LOEVVYLVSQGADPDEIGLMNIDQLPVLEYVQPGLDIIKELTSPRLIKSHLPYRFLPSD 120
QY 121 LHNGDSKVIYMARNPKDLVSVYQFHRSLRTWTSYRGTFQECRRPMNDKLGVSWEHVQ 180
Db 121 LHNGDSKVIYMARNPKDLVSVYQFHRSLRTWTSYRGTFQECRRPMNDKLGVSWEHVQ 180
QY 181 EFWEHRMDSNVFLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAEALTECHQLVDQCCNA 240
Db 181 EFWEHRMDSNVFLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAEALTECHQLVDQCCNA 240
QY 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284
RESULT 5
ST4A1_RAT STANDARD; PRT; 284 AA.
AC P63047; O88872; Q91XS5; Q9CWF7; Q9DC97;
DT 31-AUG-2004, integrated into UniProtKB/Swiss-Prot.
DT 31-AUG-2004, sequence version 1.
DT 07-FEB-2006, entry version 14.
DE Sulfotransferase 4A1 (EC 2.8.2.-) (Brain sulfotransferase-like protein) (rBR-STL) (Nervous system sulfotransferase) (NST).
GN Name=Sult4a1; Synonyms=Sultx3;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA], TISSUE SPECIFICITY, DEVELOPMENTAL STAGE,
RP AND SUBCELLULAR LOCATION.
RC STRAIN=Sprague-Dawley; TISSUE=Brain;

```

```

RX MEDLINE=20164074; PubMed=10698717; DOI=10.1042/0264-6021.3460857;
RA Falany C.N., Xie X., Wang J., Ferrer J., Falany J.L.;
RT "Molecular cloning and expression of novel sulphotransferase-like
RT cDNAs from human and rat brain.";
RL Biochem. J. 346:857-864 (2000).
RN [2]
RN NUCLEOTIDE SEQUENCE [MRNA] OF 125-284.
RP TISSUE=Brain;
RC Martin S.C., Farb D.H.;
RT "Molecular identification of a rat nervous system cytoplasmic
RT sulfotransferase, NST.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBSJ databases.
CC -!- FUNCTION: May catalyze the sulfate conjugation of many drugs,
CC xenobiotic compounds, hormones, and neurotransmitters. Displays
CC activity towards L-triiodothyronine, thyroxine, estrone, p-
CC nitrophenol, 2-naphthylamine, and 2-naphthol. May have a role in the
CC sulfation of drugs and neurotransmitters in the CNS (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasm.
CC -!- TISSUE SPECIFICITY: Expressed in the brain, not detected in the
CC liver, kidney, spleen, heart, small intestine or testis.
CC -!- DEVELOPMENTAL STAGE: Expressed at low levels in brains of 1-day
CC old animals but increase to adult levels from 7-day old animals
CC and remain at that level in adults.
CC -!- SIMILARITY: Belongs to the sulfotransferase family.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF198699; AAF61198.1; -; mRNA.
DR EMBL; AF196343; AAK64596.1; -; mRNA.
DR HSSP; P50224; 1CJM.
DR RGD; 69292; Sult4a1.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase 1; 1.
KW Lipid metabolism; Steroid metabolism; Transferase.
FT CHAIN 1 284 Sulfotransferase 4A1.
FT NP_BIND 246 254 /FTID=PRO_0000085169.
FT SEQUENCE 284 AA; 33054 MW; FCAE940F721956FC CRC64;
Query Match 98.4%; Score 1506; DB 1; Length 284;
Best Local Similarity 97.9%; Pred. No. 1.3e-117; Indels 0; Gaps 0;
Matches 278; Conservative 4; Mismatches 2;
QY 1 MAESEAETPTGPGFESKYFEFGVRLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60
Db 1 MAESEAETPTGPGFESKYFEFGVRLPPFCRGKMEETADFPVRPSDVWIVTPKSGTSL 60
QY 61 LOEVVYLVSQGADPDEIGLMNIDQLPVLEYVQPGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LOEVVYLVSQGADPDEIGLMNIDQLPVLEYVQPGLDIIKELTSPRLIKSHLPYRFLPSD 120
QY 121 LHNGDSKVIYMARNPKDLVSVYQFHRSLRTWTSYRGTFQECRRPMNDKLGVSWEHVQ 180
Db 121 LHNGDSKVIYMARNPKDLVSVYQFHRSLRTWTSYRGTFQECRRPMNDKLGVSWEHVQ 180
QY 181 EFWEHRMDSNVFLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAEALTECHQLVDQCCNA 240
Db 181 EFWEHRMDSNVFLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAEALTECHQLVDQCCNA 240
QY 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284
RESULT 6
Q3TXYS_MOUSE PRELIMINARY; PRT; 284 AA.
AC Q3TXYS;
DT 11-OCT-2005, integrated into UniProtKB/TrEMBL.
DT 11-OCT-2005, sequence version 1.
DT 07-FEB-2006, entry version 5.

```

Visual cortex cDNA, RIKEN full-length enriched library,
clone:K530039109 product:sulfotransferase family 4A, member 1, full
insert:Seulf4a1;
GN Name=Seulf4a1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=C57BL/6J; TISSUE=Visual cortex;
RC MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RX Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Methods Enzymol. 303:19-44(1999).
[2]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=C57BL/6J; TISSUE=Visual cortex;
RC PubMed16141072; DOI=10.1126/science.1112014;
RX Carninci P., Kasukawa T., Katayama S., Gough J., Hrtlich J., Watanabe T.,
Oyama R., Ravasi S., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,
Davis M.J., Wilming L.G., Aidinis V., Allen J.E.,
Ambesi-Impombato A., Aoweller R., Aturaliya R.N., Bailey T.L.,
Bansal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
Crome M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
Di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G.,
Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,
Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,
Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
Hill D., Huminecki L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,
Kitano H., Kollis G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
Kurochkin I.V., Larau L.F., Lazarevic D., Lipovich L., Liu J.,
Liu S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
Mottagui-Tabar S., Mulder N., Nakano N., Nakachi H., Ng P.,
Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavese G., Pesole G.,
Petrovsky N., Piazza S., Reed J.C., Reid J.F., Ring B.Z., Ringwald M.,
Roost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,
Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sheng Y.,
Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
Spelling S., Stupka E., Sugtara K., Sultana R., Takenaka Y., Taki K.,
Tannoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,
Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,
Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Ninomiya N.,
Nishio T., Okada M., Plessey C., Shibata K., Shiraki T., Suzuki S.,
Tagami M., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J.,
Hayashizaki Y.;
RT "The transcriptional landscape of the mammalian genome.";
RL Science 309:1559-1563(2005).
[3]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=C57BL/6J; TISSUE=Visual cortex;
RC PubMed16141073; DOI=10.1126/science.1112009;
RX RIKEN Genome Exploration Research Group, and Genome Science Group
(Genome Network Core Team) and the FANTOM Consortium;
RT "Antisense Transcription in the Mammalian Transcriptome.";
RL Science 309:1564-1566(2005).
[4]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=C57BL/6J; TISSUE=Visual cortex;
RC MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RX Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
Nikaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,

Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
Blake J.A., Bradt D., Brusic V., Chochia C., Corbani L.E., Cousins S.,
Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazier K.S.,
Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,
Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
Maglott D.R., Maitais L., Marchionni L., McKenzie L., Miki H.,
Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramchandran S.,
Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tonita M.,
Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
[5]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=C57BL/6J; TISSUE=Visual cortex;
RC MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RX Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
Aizawa K., Iwama M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
Schriml L.M., Staebli F., Suzuki R., Tonita M., Wagner L., Washio T.,
Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
Lyons P., Marchionni L., Mashima J., Mazzatelli J., Mombaerts P.,
Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsuke S.,
Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 403:685-690(2001).
[6]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=C57BL/6J; TISSUE=Visual cortex;
RC MEDLINE=20493374; PubMed=11042159; DOI=10.1101/gr.145100;
RX Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
[7]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=C57BL/6J; TISSUE=Visual cortex;
RC MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RX Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
Konno H., Akiyama J., Nishi K., Kiyosawa T., Tashiro H., Itoh M.,
Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multipillar sequencer.";


```

RL  Genome Res. 10:1757-1771 (2000).
RN  [8]
RP  NUCLEOTIDE SEQUENCE.
RC  STRAIN=C57BL/6J; TISSUE=Visual cortex;
RA  Arakawa T., Carninci P., Fukuda S., Hashizume W., Hayaashida K.,
RA  Hori F., Iida J., Imamura K., Imotani K., Itoh M., Kanagawa S.,
RA  Kawai J., Kojima M., Konno H., Murata M., Nakamura M., Ninomiya N.,
RA  Nishiyori H., Nomura K., Ohno M., Sakazume N., Sano H., Sasaki D.,
RA  Shibata K., Shiraki T., Tagami M., Tagami Y., Waki K., Watahiki A.,
RA  Muramatsu M., Hayaashizaki Y.;
RA  Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
CC  -----
CC  Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC  Distributed under the Creative Commons Attribution-NoDerivs license
CC  -----
DR  EMBL; AK159034; BAE34779.1; -; mRNA.
DR  MGI; MGI:1888971; Sult4a1.
DR  GO; GO:0005737; C:cytoplasm; TAS.
DR  GO; GO:0008146; F:sulfotransferase activity; IDA.
DR  GO; GO:0008146; F:sulfotransferase activity; RCA.
DR  GO; GO:0006790; P:sulfur metabolism; IDA.
DR  InterPro; IPR000863; Sulfotransferase.
DR  Pfam; PF00685; Sulfotransferase 1; 1.
DR  ProDom; PD001218; Sulfotransferase; 1.
KW  Transferase.
SQ  SEQUENCE 284 AA; 33054 MW; FCAE940F7219B6FC CRC64;

Query Match          98.4%; Score 1506; DB 2; Length 284;
Best Local Similarity 97.9%; Pred. No. 1.3e-117;
Matches 278; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY  1 MAESEAETSTGCEPFSEKFFEHGVLRLPPFCRGKMEIEANFVRPSDVWIVTPKSGTSL 60
Db  1 MAESEAETGTCGFESEKFFEHGVLRLPPFCRGKMEIADFFVRPSDVWIVTPKSGTSL 60
QY  61 LOEVVTVLSQGDPPDEIGLWNIDEQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db  61 LOEVVTVLSQGDPPDEIGLWNIDEQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120
QY  121 LHNGDSKVITYMARNPKDLVSVYQFHRSLRTWSYRGTFQFCRRFNMNDKLGSGWFEHVQ 180
Db  121 LHNGDSKVITYMARNPKDLVSVYQFHRSLRTWSYRGTFQFCRRFNMNDKLGSGWFEHVQ 180
QY  181 EFWEHRMDSNVFLKYVEDMHRDLVTNVEQLARFLGVSCDKAQLAEALTECHQLVQCCNA 240
Db  181 EFWEHRMDANVFLKYVEDMHRDLVTNVEQLARFLGVSCDKAQLAESLIECHQLVQCCNA 240
QY  241 EALPVGRGVRGLWKDIFTVSMNEKEDLVYKQKMGKCDLTFDFYL 284
Db  241 EALPVGRGVRGLWKDIFTVSMNEKEDLVYKQKMGKCDLTFDFYL 284

RESULT 7
QBQV8 MOUSE PRELIMINARY; PRT; 284 AA.
AC  QBQV8;
DT  01-MAR-2003, integrated into UniProtKB/TrEMBL.
DT  01-MAR-2003, sequence version 1.
DT  07-FEB-2006, entry version 20.
DE  Adult male corpora quadrigemina cDNA, RIKEN full-length enriched
DE  library, clone:B230376P15 product:sulfotransferase family 4A, member
DE  1, full insert sequence.
GN  Name=Sult4a1;
OS  Mus musculus (Mouse).
OC  Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC  Muridae; Muridae; Murinae; Mus.
OX  NCBI_TaxID=10090;
RN  [1]
RP  NUCLEOTIDE SEQUENCE.
RC  TISSUE=Corpora quadrigemina; STRAIN=C57BL/6J;
RX  MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA  Carninci P., Hayaashizaki Y.;

```

```

RT  "High-efficiency full-length cDNA cloning.";
RL  Methods Enzymol. 303:19-44 (1999).
RN  [2]
RP  NUCLEOTIDE SEQUENCE.
RC  STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
RX  PubMed=16141072; DOI=10.1126/science.1112014;
RA  Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
RA  Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,
RA  Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,
RA  Davis M.J., Wilming L.G., Aldinis V., Allen J.E.,
RA  Ambesi-Impombato A., Apweiler R., Aturaliya R.N., Bailey T.L.,
RA  Bansal K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
RA  Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
RA  di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G.,
RA  Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,
RA  Georgii-Heminger P., Gingeras T.R., Gojobori T., Green R.E.,
RA  Gusninch S., Harbers M., Hayaashi Y., Hensch T.K., Hirokawa N.,
RA  Hill D., Huminecki L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
RA  Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,
RA  Kitano H., Kollias G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
RA  Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
RA  Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
RA  Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
RA  Mottagui-Tabar S., Mulder N., Nakano N., Nakachi H., Ng P.,
RA  Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
RA  Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavese G., Pesole G.,
RA  Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
RA  Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,
RA  Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sheng Y.,
RA  Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
RA  Sperling S., Stupka E., Sugitara K., Sultana R., Takenaka Y., Taki K.,
RA  Tammoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
RA  Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,
RA  Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,
RA  Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
RA  Wahleschdt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
RA  Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
RA  Kawashima T., Kojima M., Konno S., Kondo K., Nakano K., Ninomiya N.,
RA  Nishio T., Okada M., Plessy C., Shibata K., Shiraki T., Suzuki S.,
RA  Tagami M., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J.,
RA  Hayaashizaki Y.;
RT  "The transcriptional landscape of the mammalian genome.";
RL  Science 309:1559-1563 (2005).
RN  [3]
RP  NUCLEOTIDE SEQUENCE.
RC  STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
RX  PubMed=16141073; DOI=10.1126/science.1112009;
RG  RIKEN Genome Exploration Research Group, and Genome Science Group
RG  (Genome Network Core Team) and the FANTOM Consortium;
RT  "Antisense Transcription in the Mammalian Transcriptome.";
RL  Science 309:1564-1566 (2005).
RN  [4]
RP  NUCLEOTIDE SEQUENCE.
RC  STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
RX  MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA  Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA  Nikaide I., Oeato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA  Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA  Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA  Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA  Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA  Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA  Gaasterland T., Gariboldi M., Giasi C., Godzik A., Gough J.,
RA  Grimmond S., Gusninch S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA  Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA  Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA  Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA  Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA  Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA  Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA  Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,

```


RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Vatarado R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai K., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume M., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RA "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573 (2002).
 RN [5]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yananaka I.,
 RA Saico T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischnann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690 (2001).
 RN [6]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630 (2000).
 RN [7]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
 RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771 (2000).
 RN [8]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
 RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
 RA Kato H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
 RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,

RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC -----
 DR EMBL; AK046368; BAC32692.1; -; mRNA.
 DR HSSP; P50224; 1CJM.
 DR Ensembl; ENSMUSG00000018865; Mus musculus.
 DR MGI; MGI:1888971; Sult4a1.
 DR GO; GO:0005737; C:cytoplasm; TAS.
 DR GO; GO:0008146; F:sulfotransferase activity; IDA.
 DR GO; GO:0008146; F:sulfotransferase activity; RCA.
 DR GO; GO:0006790; P:sulfur metabolism; IDA.
 DR InterPro; IPR000863; Sulfotransferase.
 DR Pfam; PF00685; Sulfotransfer_1; 1.
 DR ProDom; PD001218; Sulfotransferase; 1.
 KW Transferase.
 SQ SEQUENCE 284 AA; 32972 MW; 38D7B556F2B9391E CRC64;
 Query Match 96.9%; Score 1483; DB 2; Length 284;
 Best Local Similarity 96.5%; Pred. No. 1.1e-115;
 Matches 274; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
 QY 1 MAESEATPTSPGFEFESKYFEFHGVRLPPFCRGKMEIANNPPVRPSDVWIVYPKSGTSL 60
 DB 1 MGSEATPTGTPGFEFESKYFEFHGVRLPPFCRGKMEIANNPPVRPSDVWIVYPKSGTSL 60
 QY 61 LQEVYLVSGQADPEIGLNMNIDQLVLEYPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
 DB 61 LQEVYLVSGQADPEIGLNMNIDQLVLEYPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
 QY 121 LHNQDSKVIYARNPKDLVSYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFEHVQ 180
 DB 121 LHNQDSKVIYARNPKDLVSYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFEHVQ 180
 QY 181 EFWHRMDSNVLFKLYEDMHRDLVTWVEQLARFLGVSCDKAQLALTEHCHQLVDQCNA 240
 DB 181 EFWHRMDANVLFKLYEDMHRDLVTWVEQLARFLGVSCDKAQLALTEHCHQLVDQCNA 240
 QY 241 EALPVGGRVGLWKDIFTVSMNEXFDLVYKQMGKCDLTFDFYL 284
 DB 241 EALPVGGRVGLWKDIFTVSMNEXFDLVYKQMGKCDLTFDFYL 284
 RESULT 8
 Q68EV3_XENLA
 ID Q68EV3_XENLA PRELIMINARY; PRT; 284 AA.
 AC Q68EV3;
 DT 11-OCT-2004, integrated into UniProtKB/TrEMBL.
 DT 11-OCT-2004, sequence version 1.
 DT 07-FEB-2006, entry version 12.
 DE MGC84327 protein.
 GN Name=MGC84327;
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Eye;
 EX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.;
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 RT initiative.";
 RL Dev. Dyn. 225:384-391 (2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Eye;
 RX MEDLINE=2238257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shennen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitling M., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Eye;
RA Klein S., Gerhard D.S.;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BC080097; AAH80097.1; -; mRNA.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
SQ SEQUENCE 284 AA; 33211 MW; 9344EE54A62729D3 CRC64;

Query Match 90.5%; Score 1385; DB 2; Length 284;
Best Local Similarity 89.1%; Pred. No. 1.7e-107;
Matches 253; Conservative 18; Mismatches 13; Indels 0; Gaps 0;

QY 1 MAESEAETPTGPFESKYFEHGVRLPPFCRGKMEETANFPVRPSDWMIVTPKSGTSL 60
DB 1 MAESEAETPTGPFESKYFEHGVRLPPFCRGKMEESDFFVRENDIWIIVTPKSGTSL 60

QY 61 LOEVVYLSQAGDPDEIGLMNIDQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120
DB 61 LOEVVYLSQAGDPDEIGLMNIDQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120

QY 121 LHNGDSKIYMARNPDLVSVYQFHRSLRTWSYRGTFQEFCCRPMNDKLGYSWFHVQ 180
DB 121 LHNGDSKIYMARNPDLVSVYQFHRSLRTWSYRGTFQEFCCRPMNDKLGYSWFHVQ 180

QY 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAEALTECHOLVQCCNA 240
DB 181 EFWDHRLDSNVLFKYEDMHRDLVTWVEQLARFLGVSVYDKAQLAEALTECHOLVQCCNA 240

QY 241 EALPVGGRVGLWKDIFTVSMNEKFDLVYQKMGKCDLTFDFYL 284
DB 241 EALPIGRGRVGLWKDIFTVSMNEKFDLVYQKMGKLDLAFENL 284

RESULT 9
ID Q52KW3_XENLA PRELIMINARY; PRT; 284 AA.
AC Q52KW3;
DT 24-MAY-2005, integrated into UniProtKB/TrEMBL.
DT 24-MAY-2005, sequence version 1.
DE 24-FEB-2006, entry version 4.
DE Hypothetical protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;

RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Eye;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RL Dev. Dyn. 225:384-391 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Eye;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitling M., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Eye;
RA Klein S., Gerhard D.S.;
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BC094166; AAH94166.1; -; mRNA.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Hypothetical protein.
SQ SEQUENCE 284 AA; 33242 MW; FFA81B7E9A8BEAA9 CRC64;

Query Match 90.3%; Score 1382; DB 2; Length 284;
Best Local Similarity 89.1%; Pred. No. 3.1e-107;
Matches 253; Conservative 18; Mismatches 13; Indels 0; Gaps 0;

QY 1 MAESEAETPTGPFESKYFEHGVRLPPFCRGKMEETANFPVRPSDWMIVTPKSGTSL 60
DB 1 MAESEAETPTGPFESKYFEHGVRLPPFCRGKMEESDFFVRENDIWIIVTPKSGTSL 60

QY 61 LOEVVYLSQAGDPDEIGLMNIDQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120
DB 61 LOEVVYLSQAGDPDEIGLMNIDQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120

QY 121 LHNGDSKIYMARNPDLVSVYQFHRSLRTWSYRGTFQEFCCRPMNDKLGYSWFHVQ 180
DB 121 LHNGDSKIYMARNPDLVSVYQFHRSLRTWSYRGTFQEFCCRPMNDKLGYSWFHVQ 180

QY 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAEALTECHOLVQCCNA 240
DB 181 EFWDHRLDSNVLFKYEDMHRDLVTWVEQLARFLGVSVYDKAQLAEALTECHOLVQCCNA 240

QY 241 EALPVGGRVGLWKDIFTVSMNEKFDLVYQKMGKCDLTFDFYL 284
DB 241 EALPIGRGRVGLWKDIFTVSMNEKFDLVYQKMGKSDLTFFBSL 284

```

RESULT 10
Q4SRF8_TETNG
ID Q4SRF8_TETNG PRELIMINARY; PRT; 284 AA.
AC Q4SRF8;
DT 19-JUL-2005, integrated into UniProtKB/TrEMBL.
DT 19-JUL-2005, sequence version 1.
DT 07-FEB-2006, entry version 4.
DE Chromosome undetermined SCAR14526, whole genome shotgun sequence.
GN ORFNames=GSTENG00013931001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15496914; DOI=10.1038/nature03025;
RA Jaillon O., Aury J.-M., Brunet F., Petit J.-L., Stange-Thomann N.,
RA Mauceli E., Bounneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicoud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salancoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Bionnot C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.-P., Couzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.-N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissenbach J., Roest Croallius H.,
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBSJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBSJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; CAAB01014526; CAF96774.1; -; Genomic DNA.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransfer 1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
SQ SEQUENCE 284 AA; 32894 MW; 414AB27E7D8DF9A6 CRC64;

Query Match 85.8%; Score 1313; DB 2; Length 284;
Best Local Similarity 85.6%; Pred. No. 1.8e-101;
Matches 243; Conservative 21; Mismatches 20; Indels 0; Gaps 0;

QY 1 MAESAETPTSTGFESEKYFEHGVRLPPFCRGKMEETANFPVRPSDWIVTYPKSGTSL 60
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 MAAREADTPSTIEPESKYFEFDGVRLLPPFCRGKMEETAGFSLRSDIWIIVTYPKSGTSL 60

QY 61 LOEVVYLSQGDADPEIGLWNIDEQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
61 LOEIVYLSQGDADPEIGLWNIDEQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPTA 120

QY 121 LHNGDSKIYMARNPDKLVVSYIQFHRSLRTWSYRGTFQFCRRFMNDKLGYSWFHVQ 180
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
121 LHNGEAKVIYMARNPDKLVVSYIQFHRSLRTWSYRGTFQFCRRFMNDKLGYSWFHVQ 180

QY 161 EFWEHRMDSNVLFLKYEDMHRDLVTWEOQLARFLGVSCDKAQLEALTEHCHOLVQCCNA 240
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
161 EFWEHRMDSNVLFLKYEDMYKDLGTFVEELARFLGVSCDKAQLEGLVSCNQLIEQCNS 240

```

```

QY 241 EALPVGRGRVGLWKDIPTVSMNEKFDLVYKQKMGKCDLTFDFYL 284
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
241 EALSVCRGRVGLWKDVFTVSNNDKFDIYRQKMKCKSLTFDFGL 284

RESULT 11
Q5RCA8_PONPY
ID Q5RCA8_PONPY PRELIMINARY; PRT; 218 AA.
AC Q5RCA8;
DT 21-DEC-2004, integrated into UniProtKB/TrEMBL.
DT 21-DEC-2004, sequence version 1.
DT 07-FEB-2006, entry version 7.
DE Hypothetical protein DKFZp45910535.
GN Name=DKFZp45910535;
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Cortex;
RG The German cDNA Consortium;
RA Koehrer K., Beyer A., Mewes H.W., Weill B., Amid C., Osanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (NOV-2004) to the EMBL/GenBank/DBSJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; CR858370; CAH90599.1; -; mRNA.
DR SMR; OSRCA8; 35-218.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransfer 1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Hypothetical protein.
SQ SEQUENCE 218 AA; 25328 MW; 87B9496A82B7B878 CRC64;

Query Match 65.0%; Score 995; DB 2; Length 218;
Best Local Similarity 95.9%; Pred. No. 5.7e-75;
Matches 185; Conservative 0; Mismatches 2; Indels 6; Gaps 1;

QY 92 POPGLDIIKELTSPRLIKSHLPYRFLPSDLHNGSKVIYMARNPDKLVVSYQFHRSLRT 151
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
32 PSP-----ELTSPRLIKSHLPYRFPSPDLHNGSKVIYMARNPDKLVVSYQFHRSLRT 85

QY 152 MSYRGTFQFCRRFMNDKLGYSWFHVQEFWEHRMDSNVLFLKYEDMHRDLVTWEOQLA 211
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
86 MSYRGTFQFCRRFMNDKLGYSWFHVQEFWEHRMDSNVLFLKYEDMHRDLVTWEOQLA 145

QY 212 RFLGVSCDKAQLEALTEHCHOLVQCCNAEALPVGRGRVGLWKDIFTVSMNEKFDLVYKQ 271
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
146 RFLGVSCDKAQLEALTEHCHOLVQCCNAEALPVGRGRVGLWKDIFTVSMNEKFDLVYKQ 205

QY 272 KMGKCDLTFDFYL 284
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
206 KMGKCDLTFDFYL 218

RESULT 12
Q9R1S5_MOUSE
ID Q9R1S5_MOUSE PRELIMINARY; PRT; 298 AA.
AC Q9R1S5;
DT 01-MAY-2000, integrated into UniProtKB/TrEMBL.
DT 01-MAY-2000, sequence version 1.
DT 07-FEB-2006, entry version 18.
DE Arylsulfotransferase STIA4.
GN Name=Sult1a1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.

```


RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning.";
 RL Methods Enzymol. 303:19-44(1999).
 [12]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Gall bladder;
 RX PubMed16141072; DOI=10.1126/science.1112014;
 RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
 RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,
 RA Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,
 RA Davis M.J., Wilming L.G., Allen J.E.,
 RA Ambesi-Impombato A., Aweiller R., Aturaliya R.N., Bailey T.L.,
 RA Bansal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
 RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
 RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
 RA Di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G.,
 RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,
 RA Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,
 RA Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
 RA Hill D., Hummel L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
 RA Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,
 RA Kitano H., Kollias G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
 RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
 RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
 RA Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
 RA Mottagui-Fabriz S., Mulder N., Nakano N., Nakachi H., Ng P.,
 RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
 RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavese G., Pesole G.,
 RA Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
 RA Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,
 RA Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sheng Y.,
 RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
 RA Sperling S., Stupka E., Sugtara K., Sultana R., Takenaka Y., Taki K.,
 RA Tamoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
 RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,
 RA Yamanihi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,
 RA Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
 RA Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
 RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
 RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
 RA Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Ninomiya N.,
 RA Nishio T., Okada M., Plessy C., Shibata K., Shiraki T., Suzuki S.,
 RA Tagami M., Waki K., Watahiki A., Okamura-Ono Y., Suzuki H., Kawai J.,
 RA Hayashizaki Y.;
 RT "The transcriptional landscape of the mammalian genome.";
 RL Science 309:1559-1563(2005).
 [13]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Gall bladder;
 RX PubMed16141073; DOI=10.1126/science.1112009;
 RG RIKEN Genome Exploration Research Group, and Genome Science Group
 (Genome Network Core Team) and the FANTOM Consortium;
 RT "Antisense Transcription in the Mammalian Transcriptome";
 RL Science 309:1564-1566(2005).
 [14]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Gall bladder;
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nikaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
 RA Baldarelli R., Hall D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Baldarelli R., Hall D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Gough J.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Nunata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,

RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirazane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Alzawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 [15]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Gall bladder;
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Alzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 [16]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Gall bladder;
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayatsu M., Sugahara Y., Shibata K., Itoh M.,
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630(2000).
 [17]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Gall bladder;
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Alzawa K., Nagaoka S., Sasaki N., Carninci P.,
 RA Konno H., Akiyama J., Nishi K., Kiteunai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multiplexed sequencer.";
 RL Genome Res. 10:1757-1771(2000).
 [18]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Gall bladder;
 RA Arakawa T., Carninci P., Fukuda S., Hashizume W., Hayashida K.,
 RA Hori F., Iida J., Imamura K., Imotani K., Itoh M., Kanagawa S.,
 RA Kawai J., Kojima M., Konno H., Murata M., Nakamura M., Ninomiya N.,
 RA Nishiyori H., Nomura K., Ohno M., Sakazume N., Sano H., Sasaki D.,
 RA Shibata K., Shiraki T., Tagami M., Tagami Y., Waki K., Watahiki A.,
 RA Muramatsu M., Hayashizaki Y.;
 RL Submitted (MAR-2004) to the ENBL/GenBank/DBJ databases.
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>

"Complete sequencing and characterization of 21,243 full-length human cDNAs.";
Nat. Genet. 36:40-45 (2004).
[10]
NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schaefer C.F., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.P., Jordan K., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Frange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S.S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
"Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[11]
NUCLEOTIDE SEQUENCE OF 139-198.
RX TISSUE=Lymphocyte; PubMed=7829089;
RX MEDLINE=95130098; PubMed=7829089;
RA Aksoy I.A., Callen D.F., Apostolou S., Her C., Weinschilboum R.M.;
RT "Thermolabile phenol sulfotransferase gene (STM): localization to human chromosome 16p11.2.";
RL Genomics 23:275-277 (1994).
[12]
CHARACTERIZATION.
RX MEDLINE=94379981; PubMed=8093002;
RA Veronese M.E., Burgess W., Zhu X., McManus M.E.;
RT "Functional characterization of two human sulfotransferase cDNAs that encode monoamine- and phenol-sulphating forms of phenol sulfotransferase: substrate kinetics, thermal-stability and inhibitor-sensitivity studies.";
RL Biochem. J. 302:497-502 (1994).
[13]
X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
RX TISSUE=Brain;
RX MEDLINE=20013065; PubMed=10543947; DOI=10.1006/jmbi.1999.3153;
RA Bidwell L.M., McManus M.E., Gaedigk A., Kakuta Y., Negishi M., Pedersen L., Martin J.L.;
RT "Crystal structure of human catecholamine sulfotransferase.";
J. Mol. Biol. 293:521-530 (1999).
CC -!- FUNCTION: Catalyzes the sulfate conjugation of phenolic monoamines (neurotransmitters such as dopamine, norepinephrine and serotonin) and phenolic and catechol drugs.
CC -!- CATALYTIC ACTIVITY: 3'-phosphoadenylyl sulfate + a phenol = adenosine 3',5'-bisphosphate + an aryl sulfate.
CC -!- SUBUNIT: Homodimer.
CC -!- SUBCELLULAR LOCATION: Cytoplasm.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=P50224-1; Sequence=Displayed;
CC Name=2;
CC IsoId=P50224-2; Sequence=VSP_012326, VSP_012327;
CC Note=No experimental confirmation available;
CC -!- TISSUE SPECIFICITY: Liver, colon, kidney, lung, brain, spleen, small intestine, placenta and leukocyte.
CC -!- PTM: The N-terminus is blocked.
CC -!- SIMILARITY: Belongs to the sulfotransferase family.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License

CC EMBL; L19956; AAA02943.1; -; mRNA.
DR EMBL; L25275; AAA36523.1; -; mRNA.
DR EMBL; U08032; AAA17723.1; -; mRNA.
DR EMBL; U20499; AAA64490.1; -; Genomic_DNA.
DR EMBL; X84653; CAA59146.1; -; mRNA.
DR EMBL; L34160; -; NOT_ANNOTATED_CDS; Genomic_DNA.
DR EMBL; U37686; AAA86536.1; -; Genomic_DNA.
DR EMBL; U34199; AAC99987.1; -; mRNA.
DR EMBL; AK122733; BAC85507.1; -; mRNA.
DR EMBL; BC014471; AAH14471.1; -; mRNA.
DR EMBL; BC078144; AAH78144.1; -; mRNA.
DR EMBL; U08099; AAA82126.1; -; Genomic_DNA.
DR FIR; A55451; A55451.
DR PDB; 1CUM; X-ray; A=1-295.
DR PDB; 2A3R; X-ray; A/B=1-295.
DR OGP; P50224; -.
DR H-InvDB; HIX0012928; -.
DR HGNC; HGNC:11455; SULT1A3.
DR MIM; 600641; gene.
Query Match 30.8%; Score 472; DB 1; Length 295;
Best Local Similarity 35.8%; Pred. No. 5.4e-31;
Matches 100; Conservative 52; Mismatches 101; Indels 26; Gaps 4;
QY 24 GVRPPPCRGKMEIANPVPSPDWIVITYPKSGTSLLOEVYLVVSQADPDEIGLMNID 83
DB 17 GVPLIKYFAEALGPLQSFQARPDLLINTYPKSGTTWVSQILDMITYQGDLEKCNRAPIY 76
QY 84 EQLPVLVYQPQ-GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARPKDLV 139
DB 77 VRVPFLEVNDQGPSPGLSTLKDTPPRLIKSHLPALLPQTLDDQKVKVYVARNPKDVA 136
QY 140 VSYTQFHRSLRTMSYRGTFQFRCRRFMNDKLGYSWFHVEHHRMDSNVFLFKYEDM 199
DB 137 VSYHFFHMEKAHPPEPGTWSDFLEKFMAGEVSYGSGYQHVQEWELSRTHPVLYLFYEDM 196
QY 200 HRDLVTWVEQLARFLGVSCDKAQLEALTEHC-----HQLVDQCCNABA 242
DB 197 KENPKREIQKLEFVGRSLPEETDMFMVQHTSFKEMKKNPMTNTYTPQELMDHSIS--- 253
QY 243 LP-VQGRGVGLWKDIFTVSMNEKPDLYVKQKMGKCDLTF 280
DB 254 -PFMRKGMAGDWKTTFTVAQNERFDADYAEKNAGCSLSF 291
RESULT 17
STK1_RAT
ID STK1_RAT STANDARD; PRT; 296 AA.
AC Q9WUW8;
DT 01-DEC-2000, integrated into UniProtKB/Swiss-Prot.
DT 01-NOV-1999, sequence version 1.
DT 21-FEB-2006, entry version 32.
DE Sulfotransferase K1 (EC 2.8.2.-) (rSULT1C2).
GN Name=Sultki; Synonyms=Sultic2;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA].
RC TISSUE=Kidney;
RX MEDLINE=20329235; PubMed=10872834; DOI=10.1006/bbrc.2000.2744;
RA Xiangrong L., Joehnk C., Hartmann D., Schestag F., Kroemer W., Gieselmann V.;
RT "Enzymatic properties, tissue-specific expression, and lysosomal location of two highly homologous rat SULT1C2 sulfotransferases.";
RL Biochem. Biophys. Res. Commun. 272:242-250 (2000).
CC -!- FUNCTION: Catalyzes the sulfate conjugation of many drugs, xenobiotic compounds, hormones, and neurotransmitters.
CC -!- SUBCELLULAR LOCATION: Lysosome.
CC -!- TISSUE SPECIFICITY: Highly expressed in kidney and at lower levels


```
RESULT 21
Q2TAB3 HUMAN PRELIMINARY; PRT; 295 AA.
ID Q2TAB3 HUMAN PRELIMINARY; PRT; 295 AA.
AC Q2TAB3
DT 24-JAN-2006, integrated into UniProtKB/TrEMBL.
DT 24-JAN-2006, sequence version 1.
DE 07-FEB-2006, entry version 2.
DE Sulfotransferase family, cytosolic, 1A, phenol-preferring, member 3.
GN Name=SULT1A3;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Frange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
(2)
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Brain;
RG NIH MGC Project;
RL Submitted (DEC-2005) to the EMBL/GenBank/DBSJ databases.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC EMBL; BC111011; AA111012.1; -; mRNA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR Transferase.
KW SEQUENCE 295 AA; 34325 MW; ECDDEC18DBE80D56 CRC64;
Query Match 30.3%; Score 464; DB 2; Length 295;
Best Local Similarity 35.5%; Pred. No. 2.5e-30;
Matches 99; Conservative 52; Mismatches 102; Indels 26; Gaps 4;
QY 24 GVRLPFCRCGKMEIANFVRPSDVWIVTPKSGTSLLEQVYLVYVLSQADDPDEIGLMNID 83
DB 17 GVPLIKYFAEALGPLQSQFQARPDLLINTYPSKGTITWVSQILDMIYQGDLEKCNRAPIY 76
QY 84 EQLPVLEYPQP----GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARHPKDLV 139
DB 77 VRVPLEVNDPGEPSGLTKDTPPPRLIKSHLPALPQTLLDQKVYVYVARNPKDVA 136
QY 140 VSYVYQFHRSLRTMSYRGTFQFCRCRPMNDKLGYSWFHFVQFWEHRMDSNVLFKLYEDM 199
DB 137 VSYTHFHRMEKAPPEPGTWDSFLKFMAGEVSYNSWYHQVQEWELSRTHPVLYLFYEDM 196
QY 200 HRDLVTWVEQLARFLGVSCDKAQLAEALTEHC-----HQLVDQCNCABA 242
DB 197 KENPKREIQKILEFVGRSLPEETDMFVQHTSFKEMKKNPMNTYTVTPQELMDHSIS---- 253
QY 243 LP-VGRGRVGLWKDIFTVMNKEFDLVYKQKMGKCDLTF 280
DB 254 -PFMRKMGAGDWKTKTFTVAQNERFDADYAEKMGCSLSF 291
RESULT 22
Q58CV8 BOVIN PRELIMINARY; PRT; 298 AA.
ID Q58CV8 BOVIN PRELIMINARY; PRT; 298 AA.
AC Q58CV8;
DT 26-APR-2005, integrated into UniProtKB/TrEMBL.
DT 26-APR-2005, sequence version 1.
DE 07-FEB-2006, entry version 5.
DE Sulfotransferase family, cytosolic, 1C, member 2.
GN Name=SULT1C2;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Pooled;
RX MEDLINE=21180013; PubMed=11282978; DOI=10.1101/gr.170101;
RA Smith T.P.L., Grosse W.M., Freking B.A., Roberts A.J., Stone R.T.,
RA Casas E., Wray J.E., White J., Cho J., Fahrenkrug S.C., Bennett G.L.,
RA Heaton M.P., Laegreid W.W., Rohrer G.A., Chitko-McKown C.G.,
RA Perteira G., Holt I., Karamycheva S., Liang F., Quackenbush J.,
RA Keefe J.W.;
RT "Sequence evaluation of four pooled-tissue normalized bovine cDNA
RT libraries and construction of a gene index for cattle.";
RL Genome Res. 11:626-630 (2001).
(2)
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Pooled;
RA Hartay G.P., Sonstegard T.S., Van Tassel C.P., Clawson M.L.,
RA Heaton M.P., Keefe J.W., Snelling W.M., Weidmann R.T., Smith T.P.L.;
RT "Sequencing and analysis of Bos taurus full-length insert cDNA
RT clones.";
RL Submitted (MAR-2005) to the EMBL/GenBank/DBSJ databases.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC EMBL; BT021839; AAX46686.1; -; mRNA.
DR SMR; Q58CV8; 11-262.
DR Ensembl; ENSBTAG0000016032; Bos taurus.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase_1; 1.
KW Transferase.
SQ SEQUENCE 298 AA; 34914 MW; BE97E9EA1032E37E CRC64;
Query Match 30.3%; Score 463; DB 2; Length 298;
Best Local Similarity 35.3%; Pred. No. 3.1e-30;
Matches 98; Conservative 58; Mismatches 102; Indels 20; Gaps 5;
QY 24 GVRLP-PFCRCGKMEIANFVRPSDVWIVTPKSGTSLLEQVYLVYVLSQADDPDEIGLMNI 82
DB 19 GILQPKPTC-DTWDQIWSFQARPDLLISTYPKAGTTWTQEIVDLIQGGDVNQSRAPT 77
QY 83 DEQLPVLEYPQP----GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARHPKDL 138
DB 78 HERPFPIEWIITSLSGLEQANAMASPRMLKTHLPFHLPPSFLBENCMIYVAARNPKDN 137
QY 139 VSYVYQFHRSLRTMSYRGTFQFCRCRPMNDKLGYSWFHFVQFWEHRMDSNVLFKLYED 198
DB 138 MVSYYTHFHRMARNLPAGTWEEYFESFLAGKVCWSWDYDHVKGWQAKDQHRILLYFED 197
QY 199 MHRDLVTWVEQLARFLGVSCDKAQLAEALTEHCQLV---DQCCNAEALP----- 244
DB 198 MKENPKREIQKLAELFIKSLDKVLDKIVDHTFSFSVMKQNPMNTYTSIPNEYMNLISPF 257
```

```
QY 245 VGRGRVGLWKDFTVSMNEKFDLVYKQKMGKCDLTFDF 282
DB 258 MRKGVIGDWKNHFTVAQNERFDDYRKNMADTTLTLHF 295

RESULT 23
STBL1_CHICK
ID STBL1_CHICK STANDARD; PRT; 296 AA.
AC Q8JG30;
DT 04-JAN-2005, integrated into UniProtKB/Swiss-Prot.
DT 01-OCT-2002, sequence version 1.
DT 07-FEB-2006, entry version 20.
DE Sulfotransferase family cytosolic 1B member 1 (EC 2.8.2.-)
DE (Sulfotransferase 1B).
GN Name=SULT1B1; Synonyms=SULT1B;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA], AND FUNCTION.
RC TISSUE=Lymphoid;
RX PubMed=15234270; DOI=10.1016/j.abb.2004.05.008;
RA Wilson L.A., Reynolds G.E., Darras V.M., Coughtrie M.W.H.;
RT "cDNA cloning, functional expression, and characterization of chicken
sulfotransferases belonging to the SULT1B and SULT1C families.";
RL Arch. Biochem. Biophys. 428:64-72(2004).
CC -!- FUNCTION: Catalyzes the sulfate conjugation of many hormones,
neurotransmitters, drugs and xenobiotic compounds. Sulfonation
increases the water solubility of most compounds, and therefore
their renal excretion, but it can also result in bioactivation to
form active metabolites. Sulfates thyroid hormones including 3,3'-
diiodothyronine, catechols such as 4-nitrocatechol and 4-
isopropylcatechol and phenols such as 2-bromophenol.
CC -!- SUBCELLULAR LOCATION: Cytoplasm (By similarity).
CC -!- SIMILARITY: Belongs to the sulfotransferase family.
CC -----
CC Copyright by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonCommercial License
CC -----
DR EMBL; AJ494980; CAD41949.1; -; mRNA.
DR HSPB; P49888; IY33.
DR Ensembl; ENSGALG00000011812; Gallus gallus.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase 1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Lipid metabolism; Steroid metabolism; Transferase.
FT CHAIN 1 296 Sulfotransferase family cytosolic 1B
member 1.
FT FT NP_BIND 48 53 /FTID=PRO_0000085164.
FT FT NP_BIND 131 139 PAPS (By similarity).
FT FT NP_BIND 194 230 PAPS (By similarity).
FT FT ACT_SITE 109 109 PAPS (By similarity).
FT FT BINDING 259 259 Proton acceptor (By similarity).
FT SQ SEQUENCE 296 AA; 34074 MW; 487AB790416995F5 CRC64;

Query Match 30.0%; Score 459.5; DB 1; Length 296;
Best Local Similarity 35.8%; Pred. No. 6.1e-30;
Matches 96; Conservative 50; Mismatches 95; Indels 27; Gaps 4;

QY 36 EBIANFPVRPSDWIYTPKSGTSLQEVYVLVSQADPDEIGLMNIDEQLPVLVEYPOP- 94
DB 29 ERVDNFQSRPDDIVATPKSGTTSWSEIVDMIOGGDPKKCKRDAIVNRVPMLEFAAPG 88
QY 95 ----GLDIKEITSPRLIKSHLPYRPLSDLHNGDSKVITYMARNPKDLVSYIYQFHSRL 150
DB 89 QMPAGTEQLENMPSRIIKTHIPANILPKSFWDKCKMIYVGRNAKVAVSYTHFDLMNK 148
QY 151 TMSYGTGTQECRRFMDNKLGVSGFHVQEFWEHRMDSNVLFLKVEDMHRDLVTMVQL 210
DB 149 LHPHPGTWQDYLEAFMAGKVAYGWSFDHVRGYWERRQBHPILYLPEDMKEDLRREKAVK 208

QY 211 ARFLGVSCDKAQLAEALTEHC-----HQLVQCCNAEALP-VGRGRVGL 252
DB 209 AQFLGRELTEVALDAIAHHTSPAMRDNPNSTNYSVVPVSLMDHGIS----PPMRKGIIGD 264

QY 253 WKDIETVSMNEKFDLVYKQKMGKCDLTF 280
DB 265 WKNHFTVAQSAHFDDQYQAQKMGAGTDLRF 292

RESULT 24
STBL1_RAT
ID STBL1_RAT STANDARD; PRT; 291 AA.
AC P17988;
DT 01-NOV-1990, integrated into UniProtKB/Swiss-Prot.
DT 01-NOV-1990, sequence version 1.
DT 07-FEB-2006, entry version 55.
DE Sulfotransferase 1A1 (EC 2.8.2.1) (Aryl sulfotransferase) (Phenol
sulfotransferase) (PST-1) (Sulfokinase) (Aryl sulfotransferase IV)
DE (ASTIV) (Tyrosine-ester sulfotransferase) (Minoxidil
sulfotransferase).
GN Name=Sult1a1; Synonyms=Sst1a1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Murioidea; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA].
RC STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=90326537; PubMed=2374726;
RA Ozawa S., Nagata K., Gong D., Yamazoe Y., Kato R.;
RT "Nucleotide sequence of a full-length cDNA (PST-1) for aryl
sulfotransferase from rat liver.";
RL Nucleic Acids Res. 18:4001-4001(1990).
RN [2]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RC STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=94131305; PubMed=8299966; DOI=10.1016/0378-1119(93)90028-2;
RA Khan A.S., Taylor B.R., Chung K., Etheredge J., Gonzales R.,
Ringer D.P.;
RT "Genomic structure of rat liver aryl sulfotransferase IV-encoding
gene.";
RL Gene 137:321-326(1993).
RN [3]
RP NUCLEOTIDE SEQUENCE OF 7-291.
RX MEDLINE=92379794; PubMed=1511441;
RA Yerokun T., Etheredge J.L., Norton T.R., Carter H.A., Chung K.H.,
Birckbichler P.J., Ringer D.P.;
RT "Characterization of a complementary DNA for rat liver aryl
sulfotransferase IV and use in evaluating the hepatic gene transcript
levels of rats at various stages of 2-acetylaminofluorene-induced
hepatocarcinogenesis.";
RL Cancer Res. 52:4779-4786(1992).
RN [4]
RP CHARACTERIZATION.
RC TISSUE=Liver;
RX MEDLINE=93191709; PubMed=8447833;
RA Cruickshank D., Sansom L.N., Veronese M.E., Mojarrahi B.,
McManus M.E., Zhu X.;
RT "cDNA expression studies of rat liver aryl sulphotransferase.";
RL Biochem. Biophys. Res. Commun. 191:295-301(1993).
RN [5]
RP CHARACTERIZATION.
RX MEDLINE=95196680; PubMed=7889867;
RA Yamazoe Y., Ozawa S., Nagata K., Gong D.-W., Kato R.;
RT "Characterization and expression of hepatic sulfotransferase involved
in the metabolism of N-substituted aryl compounds.";
RL Environ. Health Perspect. 102:99-103(1994).
RN [6]
RP PROTEIN SEQUENCE OF 63-68, AND CHARACTERIZATION.
RX MEDLINE=95074030; PubMed=7982943;
RA Zheng Y., Bergold A., Duffel M.W.;
```


AC P50226; P78393;
 DT 01-OCT-1996, integrated into UniProtKB/Swiss-Prot.
 DT 07-FEB-2006, entry version 1.
 DE Sulfotransferase IA2 (EC 2.8.2.1) (Aryl sulfotransferase 2) (Phenol
 DE sulfotransferase 2) (Phenol-sulfating phenol sulfotransferase 2) (P-
 DE PST 2) (STIA2).
 GN Name=SULT1A2; Synonyms=STP2;
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Liver;
 RA Zhu X., Veronese M.E., Iocco P., McManus M.E.;
 RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Liver;
 RX MEDLINE=96065417; PubMed=7581483;
 RA Ozawa S., Nagata K., Shimada M., Ueda M., Tsuzuki T., Yamazoe Y.,
 RA Kato R.;
 RT "Primary structures and properties of two related forms of aryl
 RT sulfotransferases in human liver.";
 RL Pharmacogenetics 5:S135-S140(1995).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=94306556; PubMed=8033246; DOI=10.1016/0009-2797(94)90057-4;
 RA Yamazoe Y., Nagata K., Ozawa S., Kato R.;
 RT "Structural similarity and diversity of sulfotransferases.";
 RL Chem. Biol. Interact. 92:107-117(1994).
 RN [4]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=96299636; PubMed=8661000; DOI=10.1006/geno.1996.0216;
 RA Her C., Raftogianis R., Weinshilboum R.M.;
 RT "Human phenol sulfotransferase STP2 gene: molecular cloning,
 RT structural characterization, and chromosomal localization.";
 RL Genomics 33:409-420(1996).
 RN [5]
 RP NUCLEOTIDE SEQUENCE.
 RX PubMed=8912648; DOI=10.1006/bbrc.1996.1628;
 RA Dooley T.P., Huang Z.;
 RT "Genomic organization and DNA sequences of two human phenol
 RT sulfotransferase genes (STP1 and STP2) on the short arm of chromosome
 RT 16.";
 RL Biochem. Biophys. Res. Commun. 228:134-140(1996).
 RN [6]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=97237042; PubMed=9119390; DOI=10.1006/geno.1996.4575;
 RA Gaedigk A., Beatty B.G., Grant D.M.;
 RT "Cloning, structural organization, and chromosomal mapping of the
 RT human phenol sulfotransferase STP2 gene.";
 RL Genomics 40:242-246(1997).
 RN [7]
 RP CHARACTERIZATION.
 RX MEDLINE=95196680; PubMed=7889867;
 RA Yamazoe Y., Ozawa S., Nagata K., Gong D.-W., Kato R.;
 RT "Characterization and expression of hepatic sulfotransferase involved
 RT in the metabolism of N-substituted aryl compounds.";
 RL Environ. Health Perspect. 102:99-103(1994).
 RN [8]
 RP VARIANT ASN-235.
 RX MEDLINE=20222641; PubMed=10762004;
 RX DOI=10.1097/00008571-200003000-00008;
 RA Engelke C.E., Metzl W., Boeing H., Giatt H.;
 RT "Association between functional genetic polymorphisms of human
 RT sulfotransferases 1A1 and 1A2.";
 RL Pharmacogenetics 10:163-169(2000).
 CC -!- FUNCTION: Catalyzes the sulfate conjugation of catecholamines,
 CC phenolic drugs and neurotransmitters. Is also responsible for the
 CC sulfation and activation of minoxidil. Mediates the metabolic

CC activation of carcinogenic N-hydroxyarylamines to DNA binding
 CC products and could so participate as modulating factor of cancer
 CC risk.
 CC -!- CATALYTIC ACTIVITY: 3'-phosphoadenylyl sulfate + a phenol =
 CC adenosine 3',5'-bisphosphate + an aryl sulfate.
 CC -!- SUBUNIT: Homodimer (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasm.
 CC -!- SIMILARITY: Belongs to the sulfotransferase family.
 CC
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC
 CC EMBL; U28170; AB09659.1; -; mRNA.
 CC EMBL; U28169; AB09658.1; -; mRNA.
 CC EMBL; X78282; CAAS5088.1; -; mRNA.
 CC EMBL; U34804; AB09758.1; -; Genomic DNA.
 CC EMBL; U72202; AB08970.1; -; Genomic DNA.
 CC EMBL; U72196; AB08970.1; JOINED; Genomic DNA.
 CC EMBL; U72197; AB08970.1; JOINED; Genomic DNA.
 CC EMBL; U72198; AB08970.1; JOINED; Genomic DNA.
 CC EMBL; U72199; AB08970.1; JOINED; Genomic DNA.
 CC EMBL; U72200; AB08970.1; JOINED; Genomic DNA.
 CC EMBL; U72201; AB08970.1; JOINED; Genomic DNA.
 CC EMBL; U76619; AB18753.1; -; Genomic DNA.
 CC EMBL; U33886; AAC51149.1; -; Genomic DNA.
 CC PIR; G01843; G01843.
 CC PIR; JC5249; JC5249.
 CC HSSP; P50224; 1CJM.
 CC SMR; P50226; 8-295.
 CC Ensembl; ENSG00000197165; Homo sapiens.
 CC HGNC; HGNC:11454; SULT1A2.
 CC MIM; 601292; gene.
 CC GO; GO:0008146; P:sulfotransferase activity; TAS.
 CC GO; GO:0009309; P:amine biosynthesis; TAS.
 CC InterPro; IPR000863; Sulfotransferase.
 CC Pfam; PF00685; Sulfotransferase_1.
 CC ProDom; PD001218; Sulfotransferase; 1.
 CC KEGG; Catecholamine metabolism; Lipid metabolism; Polymorphism;
 KW Steroid metabolism; Transference.
 FT CHAIN 1 295
 FT Sulfotransferase IA2.
 FT /FTID=PRO_0000085128.
 FT NP_BIND 48 53
 FT PAPS (By similarity).
 FT NP_BIND 130 138
 FT PAPS (By similarity).
 FT NP_BIND 193 229
 FT PAPS (By similarity).
 FT NP_BIND 257 259
 FT PAPS (By similarity).
 FT ACT_SITE 108 108
 FT Proton acceptor (By similarity).
 FT VARIANT 7 7
 FT T -> I.
 FT /FTID=VAR_007426.
 FT NP_BIND 235 235
 FT PAPS (By similarity).
 FT NP_BIND 19 19
 FT PAPS (By similarity).
 FT NP_BIND 290 290
 FT PAPS (By similarity).
 FT CONFLICT 290 290
 FT P -> L (in Ref. 2).
 FT CONFLICT 290 290
 FT S -> T (in Ref. 6).
 SQ SEQUENCE 295 AA; 34286 MW; FAD88E348B49BESA CRC64;
 Query Match 29.8%; Score 456; DB 1; Length 295;
 Best Local Similarity 33.7%; Pred. No. 1.2e-29;
 Matches 94; Conservative 58; Mismatches 101; Indels 26; Gaps 4;
 QY 24 GVRPPPCRCGKMEIATNPPRPSDVWVITYPKSGTSLLEQVVLVSQGDAPDEIGLMNID 83
 DB 17 GVPLIKYFAEALGPLQSFQARPDLLIISTYPSGTTWVSQILDMEYQGDLEKCHRAPIF 76
 QY 84 EQLPVLEYQPF-GLDIIKELTSPRIKSHLPYRFLPSDLHNGDSKVIYMARNPOLIV 139
 DB 77 MRVPFLEKVPKIPSGMETLKNTPAPRLKTHLPALLPQLLDQKVKVYVARNAKOVA 136
 QY 140 VSYIQFHRSLRTMSYRGTFQFCRRFNDKLGYSWFEHQRWDSNVFLFKYEDM 199
 DB 137 VSYTHFYHMAKVPHPGTWESFLEKFNAGEVSYSQWYQHVQEWELSTHPTPLVLYFYEDM 196
 QY 200 HRDLVTWVEQLARFLGVSCDKAQLALETHC-----HQLVDQCCNAEA 242
 DB 197 KENPKREIQKILEFVGRSLPBETVDLMVMEHTSFEMKMKTPMTNTYTVTRREFMDHSIS--- 253


```
QY 243 LP-VGRGVGLWKDIFTYSMNEKFDLVYKQMGKCDLTF 280
Db 254 -PFMRKGMAGDKTFTTVAQNERFDADYAEKMGCSLSF 291

RESULT 27
Q6Y0X5_PIG
ID Q6Y0X5_PIG PRELIMINARY; PRT; 295 AA.
AC Q6Y0X5;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 9.
DE Phenol sulfotransferase.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
OC Sus.
OX NCBI_TaxID=9823;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RC PubMed=15014971; DOI=10.1007/s00335-002-2318-4;
RX Lin Z., Lou Y., Squires J.E.;
RA "Molecular cloning and functional analysis of porcine SULT1A1 gene and
RT its variant: a single mutation SULT1A1 causes a significant decrease
RT in sulfation activity.";
RL Mamm. Genome 15:218-226 (2004).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AY193893; AAP03092.1; -; mRNA.
DR HSPG; P49891; 1AQU.
DR SMR; Q6Y0X5; 9-295.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Transferase.
SQ SEQUENCE 295 AA; 34411 MW; 1721B635F6B94EE0 CRC64;

Query Match 29.8%; Score 456; DB 2; Length 295;
Best Local Similarity 35.6%; Pred. No. 1.2e-29;
Matches 98; Conservative 51; Mismatches 108; Indels 18; Gaps 3;

QY 24 GVRLPFCRCGMEIEANFPVRPSDVIWVTPKSGTSLQEVVYVLSQGDADPDEIGLMNID 83
Db 17 GVPLIKYFAEALGPLESQAMPDDVLISTPKSGTTWVSEILLDIYQGDLLQKQCRAPIF 76

QY 84 BOLPVLVEYQPQ----GLDIIKELTSPRIKSHLPYRFLPSDLHNGDSKVITYMARNPKDLV 139
Db 77 VRVPLEFIPRCPTGFELKDTAPAPLLKTHPLTLPTLLDQKVKVYVARNAKDVA 136

QY 140 VSYIQPHSLRTMSYRGTFQFCRCRPMNDKLGYSWFEHVEHMDSNVFLKYEDM 199
Db 137 VSYHYFIRMAKVPYNPNGTWSDFEDFMAGEVSIGSWYQHVQSWELRTHPLVLYFYEDM 196

QY 200 HRDLVTWVEQLARFLGVSQDKAQLALTEHC---HQLVDQCCNAALP-----V 245
Db 197 KENPKREIQKILEFVGRSILPEETVEDIVQHTSFQEMKNAMNTYRTLSDLLDHSISAFM 256

QY 246 GRGRVGLWKDIFTYSMNEKFDLVYKQMGKCDLTF 280
Db 257 RKGITGDNKSTPTVAQNERFEADYAEKMGACNLRF 291

RESULT 28
STK2_RAT
ID STK2_RAT
AC Q9UW09;
DT 01-DEC-2000, integrated into UniProtKB/Swiss-Prot.
```

```
DT 01-NOV-1999, sequence version 1.
DT 07-FEB-2006, entry version 29.
DE Sulfotransferase K2 (EC 2.8.2.-) (rSULT1C2A).
GN Name=Sultk2; Synonyms=Sultic2;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA].
RC TISSUE=Kidney;
RC MEDLINE=20329235; PubMed=10872834; DOI=10.1006/bbrc.2000.2744;
RA Xiangrong L., Joehnk C., Hartmann D., Schestag F., Kroemer W.,
RA Gieselmann V.;
RT "Enzymatic properties, tissue-specific expression, and lysosomal
RT location of two highly homologous rat SULT1C2 sulfotransferases.";
RL Biochem. Biophys. Res. Commun. 272:242-250 (2000).
CC -!- FUNCTION: Catalyzes the sulfate conjugation of many drugs,
CC xenobiotic compounds, hormones, and neurotransmitters.
CC -!- SUBCELLULAR LOCATION: Lysosome.
CC -!- TISSUE SPECIFICITY: Highly expressed in kidney and at lower levels
CC in stomach and liver. More specifically found in the epithelia of
CC proximal tubules of the kidney, of the bile duct, of the gastric
CC mucosa, and in hepatocytes.
CC -!- SIMILARITY: Belongs to the sulfotransferase family.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AJ238392; CAB41461.1; -; mRNA.
DR HSPG; P50224; 1CJM.
DR SMR; Q9UW09; 12-296.
DR Ensembl; ENSRNOG0000031833; Rattus norvegicus.
DR RGD; 621064; Sultic2.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Lysosome; Transferase.
FT CHAIN 1 296
FT NP_BIND 49 54 Sulfotransferase K2.
FT NP_BIND 131 139 /FTID=PRO_0000085136.
FT NP_BIND 194 230 PAPS (By similarity).
FT NP_BIND 258 260 PAPS (By similarity).
FT ACT_SITE 109 109 PAPS (By similarity).
FT SEQUENCE 296 AA; 34840 MW; A6D0C9257C10E34B CRC64;

Query Match 29.8%; Score 456; DB 1; Length 296;
Best Local Similarity 34.7%; Pred. No. 1.2e-29;
Matches 99; Conservative 56; Mismatches 108; Indels 22; Gaps 5;

QY 16 ESKYFEFHGVRLPFCRCGMEIEANFPVRPSDVIWVTPKSGTSLQEVVYVLSQGDADP 75
Db 10 QTKLKEVAGIPLRDSVDNWSQIQTFKAKPDDLICTYPKSGTTWVSEILLDIYQGD 69

QY 76 EIGLMNIDQOLPVLEY---POP-GLDIIKELTSPRIKSHLPYRFLPSDLHNGDSKVLYM 131
Db 70 KCQRTIIQHRPFETWARPPQPSGVVDKANAPAPRILRTHPLTQLLPSTFTNNCKLYLV 129

QY 132 ARNPKDLVVSYYQPHSLRTMSYRGTFQFCRCRPMNDKLGYSWFEHVEHMDSNV 191
Db 130 ARNAKDCWVSFYHFRMCQVLPNPTWNEYFETFGKVSQGSWFDHVKGWWEIRDYQI 189

QY 192 LFLKYEDMHRDLVTWVEQLARFLGVS-----CDKAQLEALTEHCQLVDQCCNAALP-- 244
Db 190 LFLFYEDMKRDPKREIQKVMQFMGNLDEEVVDVKIVLETSPFKMKD--NPLTFSTIPKT 247

QY 245 -----VGRGRVGLWKDIFTYSMNEKFDLVYKQMGKCDLTF 280
Db 248 IMDQISPFMRKGIQVDMKNHFTVAQNERFEADYAEKMGACNLRF 292
```


Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
[11]
RN NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RA MEDLINE=99074339; PubMed=9855620;
RX Dajani R., Hood A.M., Coughtrie M.W.;
RT "A single amino acid, Glu146, governs the substrate specificity of a
human dopamine sulfotransferase, SULT1A3.";
RL Mol. Pharmacol. 54:942-948 (1998).
[12]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA], AND VARIANT HIS-213.
RC TISSUE=Colon adenocarcinoma;
RA Shichijo S., Itoh K.;
RT "Identification of immuno-peptidomimics that are recognized by tumor-
reactive CTL generated from TIL of colon cancer patients.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
[13]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA], AND VARIANTS HIS-213 AND
VAL-223.
RA Kainane N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S.,
RX Koundinya M., Raphael J., Moreira D., Kelley T., LaBaer J., Lin Y.,
RA Phelan M., Farmer A.;
RT "Cloning of human full-length cDNAs in BD Creator(TM) system donor
vector.";
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
[14]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Placenta;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RX Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Hecichul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Maruina K., Farmer A.A., Rubin G.M., Hong L.,
RX Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RX Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny K.C., Sodergren E.J., Lu X., Gibbs R.A.,
RX Fahy J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RX Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RX Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[15]
RN NUCLEOTIDE SEQUENCE OF 1-107.
RX MEDLINE=94117022; PubMed=8288252;
RA Dooley T.P., Obermoller R.D., Leiter E.H., Chapman H.D., Falany C.N.,
RX Deog Z., Siciliano M.J.;
RT "Mapping of the phenol sulfotransferase gene (STP) to human chromosome
16p12.1-pl.2 and to mouse chromosome 7.";
RL Genomics 18:440-443 (1993).
[16]
RN CHARACTERIZATION.
RX MEDLINE=94379981; PubMed=8093002;
RA Veronesi M.E., Burgess W., Zhu X., McManus M.E.;
RT "Functional characterization of two human sulfotransferase cDNAs that
encode monoamine- and phenol-sulphating forms of phenol
sulfotransferase: substrate kinetics, thermal-stability and
inhibitor-sensitivity studies.";
RL Biochem. J. 302:497-502 (1994).
[17]
RN MUTAGENESIS OF CYS-70.
RX MEDLINE=94306582; PubMed=8033270; DOI=10.1016/0009-2797(94)90053-1;
RA Falany C.N., Zhuang W., Falany J.L.;
RT "Characterization of expressed human phenol-sulfating phenol
sulfotransferase: effect of mutating cys70 on activity and
thermostability.";
Chem. Biol. Interact. 92:57-66 (1994).
[18]
RN VARIANTS GLN-37; HIS-213 AND VAL-223.
RX MEDLINE=98005125; PubMed=9345314; DOI=10.1006/bbr.1997.7466;
RA Raftogiannis R.B., Wood T.C., Otterness D.M., Van Loon J.A.,
RX Weinshilboum R.M.;
RT "Phenol sulfotransferase pharmacogenetics in humans: association of
common SULT1A1 alleles with RS PST phenotype.";
RL Biochem. Biophys. Res. Commun. 239:298-304 (1997).
[19]
RN VARIANT HIS-213.
RX MEDLINE=20222641; PubMed=10762004;
DOI=10.1097/00008571-200003000-00008;
RA Engelte C.E., Meinel W., Boeing H., Glatt H.;
RT "Association between functional genetic polymorphisms of human
sulfotransferases 1A1 and 1A2.";
RL Pharmacogenetics 10:163-169 (2000).
[20]
RN X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) IN COMPLEX WITH
ADENOSINE-3'-5'-DIPHOSPHATE (PAP) AND A PHENOL.
RX PubMed=12471039; DOI=10.1074/jbc.M207246200;
Gamage N.U., Dugleby R.G., Barnett A.C., Tresillian M., Latham C.F.,
RX Liyou N.E., McManus M.E., Martin J.L.;
RT "Structure of a human carcinogen-converting enzyme, SULT1A1.
Structural and kinetic implications of substrate inhibition.";
RL J. Biol. Chem. 278:7655-7662 (2003).
CC -!- FUNCTION: Catalyzes the sulfate conjugation of catecholamines,
phenolic drugs and neurotransmitters. It is also responsible for the
sulfation and activation of minoxidil. Mediates the metabolic
activation of carcinogenic N-hydroxyarylamines to DNA binding
products and could so participate as modulating factor of cancer
risk.
CC -!- CATALYTIC ACTIVITY: 3'-phosphoadenylyl sulfate + a phenol =
adenosine 3',5'-diphosphate + an aryl sulfate.
CC -!- SUBUNIT: Homodimer.
CC -!- SUBCELLULAR LOCATION: Cytoplasm.
CC -!- TISSUE SPECIFICITY: Liver, lung, adrenal, brain, platelets and
skin.
CC -!- SIMILARITY: Belongs to the sulfotransferase family.
Query Match 29.5%; Score 451; DB 1; Length 295;
Best Local Similarity 33.8%; Pred. NO. 3.1e-29;
Matches 93; Conservative 57; Mismatches 107; Indels 18; Gaps 3;
QY 24 GVRPPPCRGKMEIANFPVPSDWIVTYPKSGTSLQEVVYLVVSQADPDEIGLMNID 83
DB 17 GVPLIKYFAEALGPLQSFQAPDPLLSTYPSKSGTTWVSQILDMYQGGLEKCHRAPIF 76
QY 84 EQLPVLEYQPQ---GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARPKDLV 139
DB 77 MRVPFLEKAPGIPSGMETLADTPAPRLKTHLPALLPQTLDDQKVVYVARNAKDVA 136
QY 140 VSYTQFHRSLRMTSYRGTFQFCRRFMNDKLGYSWFHEHQRMDSNVFLFKYEDM 199
DB 137 VSYTHFYHMAKVHPPEGTWDSFLEKFMVGEVSYGSWYQHVQEWELSRTHPVLFLFYEDM 196
QY 200 HRDLVTWVEQLARFLGVSCDAQAEALTEHC---HQLVDDCCNABALP-----V 245
DB 197 KENPKREIQKILEFVGRSLPDETFDMVQHTSFKEMKKNPMNTNTVTPQEFMDHSISPFM 256
QY 246 GRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
DB 257 RKGAGDWKTTFTVAQNERFDADYAEKMGCSLSF 291
RESULT 31
Q2NL71 HUMAN
ID Q2NL71 HUMAN PRELIMINARY; PRT; 295 AA.
AC Q2NL71
DT 07-FEB-2006, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, sequence version 1.
DT 07-FEB-2006, entry version 1.
DE SULT1A1 protein.

```
GN Name=SULT1A1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Colon;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altechul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schain J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RL [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Colon;
RG NIH MGC Project;
RL Submitted (DEC-2005) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BC110887; AA110888.1; -; mRNA.
SQ SEQUENCE 295 AA; 34197 MW; 60E9D5BBA9159176 CRC64;

Query Match 29.5%; Score 451; DB 2; Length 295;
Best Local Similarity 33.8%; Pred. No. 3.1e-29;
Matches 93; Conservative 57; Mismatches 107; Indels 18; Gaps 3;

Qy 24 GVRLPFCRGKMEETANFVRPSDVWIVTPKSGTSLLOEVVYLSQGADPDDELGMNID 83
Db 17 GVPLIKYFAEALGPQSQFQARPDLLISTYPSKGTWVSQILDMIYQGGDLKCHRAPIF 76

Qy 84 EQLPVLEYPOP----GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYARNPKDLV 139
Db 77 MRVPLEFKAPGIPSGMETLKDTPAPRLKTLPLALLPQLTLLDQKVVVYARNAKDVA 136

Qy 140 VSYQFHRSLRTMSVYRGTFQFCFRFMDKLGYSNFBHQFVFWHRMDSNVFLFKYEDM 199
Db 137 VSYHYFYHMAKVHPPEGTWDSFLEKFMVGEVSYGSWYQHVQEWELSRTHPVLYLFYEDM 196

Qy 200 HRDLVTWVEQLARFLGVSCDKAQLALTEHC---HQLVDOCCNAEALP-----V 245
Db 197 KENPKREIQLKLEFVGRSLPSETVDFMVQHTSFKEKMKNPMTNTVTVFQEFMDHSISPFM 256

Qy 246 GRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 257 RKGWAGDWKTTFTVAQNERFDADYAEKMGAGCSLSF 291

RESULT 32
Q32C30 BOVIN
ID Q32C30 BOVIN PRELIMINARY; PRT; 295 AA.
AC Q32C30
DT 27-SEP-2005, integrated into UniProtKB/TrEMBL.
DR 27-SEP-2005, sequence version 1.
```

```
DT 07-MAR-2006, entry version 5.
DE Sulfotransferase, estrogen-preferring.
GN Name=STE;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Hereford; TISSUE=Reticulum;
RA Moore S., Alexander L., Brownstein M., Guan L., Lobo S., Meng Y.,
RA Tanaguchi M., Wang Z., Yu J., Prange C., Schreiber K., Shenmen C.,
RA Chuner L., Bala M., Barbazuk S., Barber S., Babakaliff R., Beland J.,
RA Wagn E., Del Rio L., Gibson S., Hanson R., Kirkpatrick R., Liu J.,
RA Matsuo C., Mayo M., Santos R.R., Stott J., Tsai M., Wong D.,
RA Siddiqui A., Holt R., Jones S.J., Marra M.A.;
RL Submitted (AUG-2005) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BC102939; AA102940.1; -; mRNA.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase_1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Transferase.
SQ SEQUENCE 295 AA; 34660 MW; 81456185E090F016 CRC64;

Query Match 29.5%; Score 451; DB 2; Length 295;
Best Local Similarity 34.6%; Pred. No. 3.1e-29;
Matches 101; Conservative 57; Mismatches 110; Indels 24; Gaps 5;

Qy 10 STPGEFFESKYFEFGVRLPPECRGKWEIARN---PPVPSDVWIVTPKSGTSLLOEVVY 66
Db 3 SKSPFSFYFKLGGI---PMYKPIEQFHVVEFARPDDLVIIVTPKSGTTLWSEIIC 59

Qy 67 LVSGQADPDDELGMNIDQLPVL---EYPOPGLDIIKELTSPRLIKSHLPYRFLPSDLH 122
Db 60 MIYNGDVEKCKEDVIFNRPVYLECSTEHWVNGVQLNEMASPRIVKSHLPVKLLPVSPW 119

Qy 123 NGDSKVIYARNPKDLVVSYYOFHRSRLTMSYRGTFQFCFRFMDKLGYSNFBHQF 182
Db 120 EKNCCKIILSRNAKDVVVSYYFFILMTVAIPDPDSFQDFVEKFMDEVPYGSWFHTKSW 179

Qy 183 WEHRMDSNVFLKYEDMHRDLVTWVEQLARFLGVSCDKAQLALTEHC----- 231
Db 180 WEKSKNPQVLFYEDMKENIRKEYMKLLEFLGRKASDELVDKIHKHTSFQEMKNPSTN 239

Qy 232 --QLVDOCCNAEALP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 240 YTTLPDEVNMQKVSFFMRKGDVGDWKHFTVALNEKFDHMYEQQMGKSTLKF 291

RESULT 33
ST1A1 CANFA
ID ST1A1 CANFA STANDARD; PRT; 295 AA.
AC Q29476;
DT 26-APR-2005, integrated into UniProtKB/Swiss-Prot.
DT 01-NOV-1996, sequence version 1.
DT 07-FEB-2006, entry version 29.
DE Sulfotransferase 1A1 (EC 2.8.2.1) (Aryl sulfotransferase) (Phenol
DE sulfotransferase) (Phenol-sulfating phenol sulfotransferase) (P-PST).
GN Name=SULT1A1;
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
OC Canis.
OX NCBI_TaxID=9615;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA].
```

```

RC STRAIN=Beagle; TISSUE=Liver;
RA Saitakawa M., Ogura K., Nakamura T., Watabe T.;
RT "Molecular cloning and sequencing of a dog liver cDNA (dPST-1)
RT encoding a phenol sulfotransferase.";
RN Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.
RP NUCLEOTIDE SEQUENCE (MRNA), CHARACTERIZATION, AND TISSUE SPECIFICITY.
RC TISSUE=Liver;
RX MEDLINE=22049953; PubMed=12054462; DOI=10.1016/S0003-9861(02)00021-8;
RA Teo C., Morgenstern R., Swedmark S.;
RT "Canine sulfotransferase SULT1A1: molecular cloning, expression, and
RT characterization.";
RL Arch. Biochem. Biophys. 401:125-133 (2002).
CC -!- FUNCTION: Catalyzes the sulfate conjugation of catecholamines and
CC phenolic drugs and neurotransmitters. Acts on p-nitrophenol (PNP),
CC alpha-naphthol, dopamine, minoxidil and beta-estradiol but not
CC dehydroepiandrosterone.
CC -!- CATALYTIC ACTIVITY: 3'-phosphoadenylyl sulfate + a phenol =
CC adenosine 3', 5'-bisphosphate + an aryl sulfate.
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasm (By similarity).
CC -!- TISSUE SPECIFICITY: Ubiquitously expressed in canine tissues with
CC highest expression in male and female liver.
CC -!- SIMILARITY: Belongs to the sulfotransferase family.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
DR EMBL; D29807; BAA06190.1; -; mRNA.
DR EMBL; AY069922; AAL5717.1; -; mRNA.
DR HSP; P50224; 1CJM.
DR SMR; Q29476; 8-295.
DR Ensembl; ENSCAFG00000017122; Canis familiaris.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase 1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
DR Catecholamine metabolism; Lipid metabolism; Steroid metabolism;
KW Transferase.
FT CHAIN 1 295 Sulfotransferase 1A1.
FT NP_BIND 48 53 /FTID=PRO_0000085126.
FT NP_BIND 130 138 PAPS (By similarity).
FT NP_BIND 193 229 PAPS (By similarity).
FT NP_BIND 257 259 PAPS (By similarity).
FT ACT_SITE 108 108 Proton acceptor (By similarity).
SQ SEQUENCE 295 AA; 34115 MW; C297A9211A5609B6 CRC64;

Query Match 29.4%; Score 450; DB 1; Length 295;
Best Local Similarity 33.5%; Pred. No. 3.8e-29;
Matches 95; Conservative 61; Mismatches 102; Indels 26; Gaps 4;

QY 18 KYFPHGVRLPPFCRGKMEETANFVRPSDWMIVTPKSGTSLLOEVVYLVSGQADPDEI 77
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 78 GLMNTDEQLPVLEYQPQ---GLDIIKELTSPRLKSHLPYFLPSDLHNGDSKVIYMAR 133
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 134 NPKDLWSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVSFWEHRMDSNVLF 193
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 194 LKYEDEMRLVTMVQSLARFLGVSCDKAQLEALTEHCH-----QLVDQ 236
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 237 CCNAALPVGRGVGLWKIDFTVSNMKEFDLVYKQKMGKDLTF 280
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 251 SISAP---MRKGISGDWKTFTTVAQNERFDADYAKKMGCGUSP 291
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 34

```

```

Q95KM3 RABIT
ID Q95KM3_RABIT PRELIMINARY; PRT; 295 AA.
AC Q95KM3;
DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT 01-DEC-2001, sequence version 1.
DT 07-FEB-2006, entry version 16.
DE Cytosolic phenol sulfotransferase SULT1A1.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Lagomorpha; Leporidae;
OC Oryctolagus.
OX NCBI TaxID=9986;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=22003852; PubMed=12007634; DOI=10.1016/S1357-2725(02)00017-1;
RA Riley B., Bolton-Grob R., Liyou N., Wong C., Tresillian M.,
RA McManus M.E.;
RT "Isolation and characterisation of a novel rabbit sulfotransferase
RT isoform belonging to the SULT1A subfamily.";
RL Int. J. Biochem. Cell Biol. 34:958-969(2002).
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
DR EMBL; AF360872; AAK50763.1; -; mRNA.
DR HSP; P50224; 1CJM.
DR SMR; Q95KM3; 8-295.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase 1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Transferase.
SQ SEQUENCE 295 AA; 34193 MW; 78F688CD2511E484 CRC64;

Query Match 29.3%; Score 449; DB 2; Length 295;
Best Local Similarity 35.1%; Pred. No. 4.6e-29;
Matches 97; Conservative 51; Mismatches 108; Indels 20; Gaps 4;

QY 24 GVRLLPPFCRGKMEETANFVRPSDWMIVTPKSGTSLLOEVVYLVSGQADPDEI 83
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 84 EQLPVLEYQPQ---GLDIIKELTSPRLKSHLPYFLPSDLHNGDSKVIYMARPKDLV 139
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 140 VSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVSFWEHRMDSNVLF 199
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 200 HRDLVTMVQSLARFLGVSCDKAQLEALTEHCHQLVDQ---CCNAALP----- 244
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 245 VGRGVGLWKIDFTVSNMKEFDLVYKQKMGKDLTF 280
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 256 MRRGVAGDWKTFTTVAQHEFYADYAKKMGCELTF 291
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 35
ID STIC1_HUMAN STANDARD; PRT; 296 AA.
AC O00338;
DT 01-DEC-2000, integrated into UniProtKB/Swiss-Prot.
DT 01-JUL-1997, sequence version 1.
DT 07-MAR-2006, entry version 52.
DE Sulfotransferase 1C1 (EC 2.8.2.-) (SULT1C#1) (ST1C2) (humSULTC2).
GN Name=SULT1C1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;

```

OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Fetal liver, and Fetal spleen;
 RX MEDLINE=97312707; PubMed=9169148; DOI=10.1006/geno.1997.4683;
 RA Her C., Kaur G.P., Ahtwal R.S., Weinshilboum R.M.;
 RT "Human sulfoltransferase SULT1C1: cDNA cloning, tissue-specific
 expression, and chromosomal localization.";
 RL Genomics 41:467-470(1997).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Fetal liver;
 RX MEDLINE=98297811; PubMed=9635888; DOI=10.1093/carcin/19.5.951;
 RA Yoshinari K., Nagata K., Shimada M., Yamazoe Y.;
 RT "Molecular characterization of SULT1C1-related human sulfoltransferase.";
 RL Carcinogenesis 19:951-953(1998).
 RN [3]
 RP NUCLEOTIDE SEQUENCE, AND CHARACTERIZATION.
 RC TISSUE=Fetal lung;
 RX MEDLINE=99069375; PubMed=9852044; DOI=10.1074/jbc.273.51.33929;
 RA Sakakibara Y., Yanagisawa K., Katafuchi J., Ringer D.P., Takami Y.,
 RA Nakayama T., Suiko M., Liu M.-C.;
 RT "Molecular cloning, expression, and characterization of novel human
 SULT1C sulfoltransferases that catalyze the sulfonation of N-hydroxy-2-
 acetylaminofluorene.";
 RL J. Biol. Chem. 273:33929-33935(1998).
 RN [4]
 RP NUCLEOTIDE SEQUENCE, AND CHARACTERIZATION.
 RC TISSUE=Stomach;
 RX MEDLINE=99410886; PubMed=10481272; DOI=10.1016/S1357-2725(99)00038-2;
 RA Hehonah N., Zhu X., Brix L., Bolton-Grob R., Barnett A., Windmill K.,
 RA McManus M.;
 RT "Molecular cloning, expression, localisation and functional
 characterisation of a rabbit SULT1C2 sulfoltransferase.";
 RL Int. J. Biochem. Cell Biol. 31:869-882(1999).
 RN [5]
 RP NUCLEOTIDE SEQUENCE, AND ALTERNATIVE SPLICING.
 RX MEDLINE=20247255; PubMed=10783263; DOI=10.1006/geno.2000.6150;
 RA Freimuth R.R., Raftogiannis R.B., Wood T.C., Moon E., Kim U.-J., Xu J.,
 RA Siciliano M.J., Weinshilboum R.M.;
 RT "Human sulfoltransferases SULT1C1 and SULT1C2: cDNA characterization,
 gene cloning, and chromosomal localization.";
 RL Genomics 65:157-165(2000).
 RN [6]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM SHORT).
 RA Kalnine N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S.,
 RA Koundinya M., Raphael J., Moreira D., Kelley T., LaBaer J., Lin Y.,
 RA Phelan M., Farmer A.;
 RT "Cloning of human full-length CDSs in BD Creator (TM) system donor
 vector.";
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RC TISSUE=Kidney;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ussin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S.G., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny N.B., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting R.W., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -!- FUNCTION: Catalyzes the sulfate conjugation of many drugs,
 CC xenobiotic compounds, hormones, and neurotransmitters. May be
 CC involved in the activation of carcinogenic hydroxylamines. Shows
 CC activity towards p-nitrophenol and N-hydroxy-2-acetylami-
 CC fluorene (N-OH-2AAF).
 CC -!- SUBCELLULAR LOCATION: Cytoplasm.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=Short; Sequence=Displayed;
 CC IsoId=000338-1; Sequence=Displayed;
 CC Name=Long;
 CC IsoId=000338-2; Sequence=VSP_006303;
 CC -!- TISSUE SPECIFICITY: Found in adult stomach, kidney and thyroid
 CC gland, and in fetal kidney and liver.
 CC -!- SIMILARITY: Belongs to the sulfoltransferase family.
 CC -----
 CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC -----
 CC EMBL: U6036; AAC51285.1; -; mRNA.
 DR EMBL: AB008164; BAA28346.1; -; mRNA.
 DR EMBL: AF026303; AAC00409.1; -; mRNA.
 DR EMBL: AF186251; AAF72799.1; -; mRNA.
 DR EMBL: AF186252; AAF72800.1; -; mRNA.
 DR EMBL: AF186253; AAF72801.1; -; mRNA.
 DR EMBL: AF186254; AAF72802.1; -; mRNA.
 DR EMBL: AF186255; AAF72803.1; -; mRNA.
 DR EMBL: AF186256; AAF72804.1; -; mRNA.
 DR EMBL: AF186262; AAF72805.1; -; Genomic DNA.
 DR EMBL: AF186258; AAF72805.1; JOINED; Genomic DNA.
 DR EMBL: AF186260; AAF72805.1; JOINED; Genomic DNA.
 DR EMBL: AF186261; AAF72805.1; JOINED; Genomic DNA.
 DR EMBL: AF186262; AAF72806.1; -; Genomic DNA.
 DR EMBL: AF186258; AAF72806.1; JOINED; Genomic DNA.
 DR EMBL: AF186259; AAF72806.1; JOINED; Genomic DNA.
 DR EMBL: AF186260; AAF72806.1; JOINED; Genomic DNA.
 DR EMBL: AF186261; AAF72806.1; JOINED; Genomic DNA.
 DR EMBL: BT006951; AAP35597.1; -; mRNA.
 DR EMBL: BC005353; AAH05353.1; -; mRNA.
 DR PDB: 2ETG; X-ray; A/B=1-296.
 DR Ensemble; ENSG00000198203; Homo sapiens.
 DR H-InvDB; HIX0002353; -.
 DR HGNC; HGNC:11456; SULT1C1.
 DR MIM; 602385; gene.
 DR GO; GO:0005737; C:cytoplasm; TAS.
 DR GO; GO:0008146; F:sulfoltransferase activity; TAS.
 DR GO; GO:0009308; P:amine metabolism; TAS.
 DR InterPro; IPR000863; Sulfoltransferase.
 DR Pfam; PF00685; Sulfoltransferase 1; 1.
 DR ProDom; PD001218; Sulfoltransferase; 1.
 DR 3D-structure; Alternative splicing; Polymorphism; Transferase.
 KW CHAIN 1 296 Sulfoltransferase ICI.
 FT /FTId=PRO_0000085132.
 FT NP_BIND 49 54 PAPS (By similarity).
 FT NP_BIND 131 139 PAPS (By similarity).
 FT NP_BIND 194 230 PAPS (By similarity).
 FT NP_BIND 258 260 PAPS (By similarity).
 FT ACT_SITE 109 109 Proton acceptor (By similarity).
 FT VARSPLIT 93 113 SSNPASTPSQAKITD (in isoform Long).
 FT /FTId=VSP_006303.
 FT Y -> H (in dbSNP:17036091).
 FT /FTId=VAR_021986.
 FT S -> A (in dbSNP:17036104).
 FT /FTId=VAR_021987.
 FT VARIANT 128 128
 FT VARIANT 255 255
 FT STRAND 14 16
 FT TURN 17 18
 FT STRAND 19 21
 FT HELIX 23 27
 FT TURN 28 28


```

FT          29          33          34          34          34          34
FT          34          34          34          34          34          34
FT          39          40          40          40          40          40
FT          42          46          46          46          46          46
FT          48          49          49          49          49          49
FT          50          51          51          51          51          51
FT          52          64          64          64          64          64
FT          65          65          65          65          65          65
FT          82          82          82          82          82          82
FT          83          84          84          84          84          84
FT          88          88          88          88          88          88
FT          90          90          90          90          90          90
FT          93          93          93          93          93          93
FT          100         100         100         100         100         100
FT          102         103         103         103         103         103
FT          105         108         108         108         108         108
FT          112         114         114         114         114         114
FT          115         115         115         115         115         115
FT          117         117         117         117         117         117
FT          118         121         121         121         121         121
FT          122         122         122         122         122         122
FT          123         123         123         123         123         123
FT          133         146         146         146         146         146
FT          148         149         149         149         149         149
FT          150         150         150         150         150         150
FT          155         155         155         155         155         155
FT          164         164         164         164         164         164
FT          165         166         166         166         166         166
FT          168         168         168         168         168         168
FT          170         170         170         170         170         170
FT          173         183         183         183         183         183
FT          184         186         186         186         186         186
FT          187         193         193         193         193         193
FT          199         199         199         199         199         199
FT          201         211         211         211         211         211
FT          212         213         213         213         213         213
FT          218         227         227         227         227         227
FT          228         228         228         228         228         228
FT          229         229         229         229         229         229
FT          230         235         235         235         235         235
FT          262         262         262         262         262         262
FT          264         264         264         264         264         264
FT          265         268         268         268         268         268
FT          271         285         285         285         285         285

Query Match          29.3%; Score 449; DB 1; Length 296;
Best Local Similarity 33.1%; Pred. No. 4.6e-29;
Matches 94; Conservative 61; Mismatches 105; Indels 24; Gaps 4;

QY 18 KYFPHGVRLPPFCRGKMEETANFPVPSDVWIVTPKSGTSLLOEVVLYVLSQGADPDEI 77
DB 12 KLKEVEGTLQPATVDNNWSQISQSFAPKPPDDLLICTYPKAGTTWQIEIVDMIEQNGDVKEC 71

QY 78 GLMNIDEOLPVLEY---POP-GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMAR 133
DB 72 QRALIQHRHPTEIWARPPQPSGVEKAKAMPSPRIKTHLSQLLPPSPWNNCKFLYVAR 131

QY 134 NPKDLVSVYQPHRSRLRTMSYRGTPQFCRPFMDKLGYSWFHQBFWFHRMDSNYLF 193
DB 132 NAKDCMVSYYHFORNMHMLPDGPTWEEYFETFKNGKVVWGSFWDHVKGWEMKDRHQILF 191

QY 194 LKYEDMHRLVTWVQLARFLGVSCDKQAQLEALTEH-----CHQLVYDQ 236
DB 192 LFYEDIKRDPKHEIRKVMQFMGKKYDQETVLDKIVQETSFEMKENPMNTRSTVSKSLDQ 251

QY 237 CCNAELPVGRVGLWKDIFTVSMNEKFDLYVKQKMGKDLTF 280
DB 252 SISSF--MRKGTGVDWKNHFTVAQNERPDEIYRKMEGTSINF 292

RESULT 36
ST1C2 HUMAN
ID ST1C2_HUMAN STANDARD; PRT; 302 AA.

```

```

AC 075897; O53S63;
DT 01-DEC-2000, integrated into UniProtKB/Swiss-Prot.
DT 07-MAR-2006, sequence version 2.
DE 07-MAR-2006, entry version 37.
DE Sulfotransferase 1C2 (EC 2.8.2.-) (SULT1C) (SULT1C#2).
GN Name=SULT1C2;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA], AND VARIANT GLU-5.
RC TISSUE=Petal lung;
RX MEDLINE=99069375; PubMed=9852044; DOI=10.1074/jbc.273.51.33929;
RA Sakakibara Y., Yanagisawa K., Katafuchi J., Ringer D.P., Takami Y.,
RA Nakayama T., Suiko M., Liu M.-C.;
RT "Molecular cloning, expression, and characterization of novel human
RT SULT1C sulfotransferases that catalyze the sulfonation of N-hydroxy-2-
RT acetylaminofluorene.";
RL J. Biol. Chem. 273:33929-33935(1998).
RN [2]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RX MEDLINE=20247255; PubMed=10783263; DOI=10.1006/geno.2000.6150;
RA Preimuth R.R., Rafiogianis R.B., Wood T.C., Moon E., Kim U.-J., Xu J.,
RA Siciliano M.J., Weinsilboum R.M.;
RT "Human sulfotransferases SULT1C1 and SULT1C2: cDNA characterization,
RT gene cloning, and chromosomal localization.";
RL Genomics 65:157-165(2000).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15815621; DOI=10.1038/nature03466;
RA Hillier L.W., Graves T.A., Fulton R.S., Fulton L.A., Pepin K.H.,
RA Minx P., Wagner-McPherson C., Layman D., Wyllie K., Sekhon M.,
RA Becker M.C., Fewell G.A., Delehaunty K.D., Miner T.L., Nash W.E.,
RA Krenitzki C., Oddy L., Du H., Sun H., Bradshaw-Cordum H., Ali J.,
RA Carter J., Cordes M., Harris A., Isak A., van Brunt A., Nguyen C.,
RA Du P., Courtney L., Kalicki J., Ozersky P., Abbott S., Armstrong J.,
RA Beiter E.A., Caruso L., Cedroni M., Cotton M., Davidson T., Desai A.,
RA Elliott G., Erb T., Fronick C., Gaige T., Haakenson W., Haglund K.,
RA Holmes A., Harkins R., Kim K., Kruchowski S.S., Strong C.M.,
RA Grewal N., Goyea E., Hou S., Levy A., Martinka S., Mead K.,
RA McLellan M.D., Meyer R., Randall-Maher J., Tomlinson C.,
RA Dauphin-Kohlberg S., Kozlowski-Reilly A., Shah N.,
RA Swearingen-Shahid S., Snider J., Strong J.T., Thompson J., Yoakum M.,
RA Leonard S., Fearman C., Trani L., Radionenko M., Waligorski J.E.,
RA Wang C., Rock S.M., Tin-Wollam A.-M., Maupin R., Latreille P.,
RA Wendt M.C., Yang S.-P., Pohl C., Wallis J.W., Spieth J., Bieri T.A.,
RA Berkowicz N., Nelson J.O., Osborne J., Ding L., Meyer R., Sabo A.,
RA Eldred J., Williams D., Jones T.A., She X., Ciccarelli F.D.,
RA Shatland Y., Sinha P., Wohldmann P.E., Cook L.L., Hickenbotham M.T.,
RA Izaurrealde E., Taylor J., Schmutz J., Myers R.M., Cox D.R., Huang X.,
RA McPherson J.D., Mardis E.R., Clifton S.W., Warren W.C.,
RA Chinwalla A.T., Eddy S.R., Marra M.A., Ovcharenko I., Furey T.S.,
RA Miller W., Eichler E.E., Bork P., Suyama M., Torrents D.,
RA Waterston R.H., Wilson R.K.;
RT "Generation and annotation of the DNA sequences of human chromosomes 2
RT and 4.";
RL Nature 434:724-731(2005).
CC -!- FUNCTION: Catalyzes the sulfate conjugation of many drugs,
CC xenobiotic compounds, hormones, and neurotransmitters. May be
CC involved in the activation of carcinogenic hydroxylamines. Shows
CC activity towards p-nitrophenol and N-hydroxy-2-acetylmino-
CC fluorene (N-OH-2AAF).
CC -!- SUBCELLULAR LOCATION: Cytoplasm.
CC -!- TISSUE SPECIFICITY: Expressed at high levels in fetal lung and
CC kidney and at low levels in fetal heart, adult kidney, ovary and
CC spinal chord.
CC -!- SIMILARITY: Belongs to the sulfotransferase family.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC

```



```
DR EMBL; AF055584; AAC95519.1; -; mRNA.
DR EMBL; AF186263; AAF72810.1; -; Genomic DNA.
DR EMBL; AC068941; AAY14742.1; -; Genomic DNA.
DR PDB; 2AD1; X-ray; A=7-302.
DR Ensembl; ENSG00000198075; Homo sapiens.
DR HGNC; HGNC:11457; SULT1C2.
DR MIM; 608357; gene.
DR LinkHub; O75897; -.
DR GO; GO:0008146; F:sulfotransferase activity; NAS.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase 1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
DR 3D-structure; Polymorphism; Transferase.
KW CHAIN 1 302 Sulfotransferase 1C2.
FT FT
FT NP_BIND 55 60 /FTID=PRO 0000085137.
FT NP_BIND 137 145 PAPS (By similarity).
FT NP_BIND 200 236 PAPS (By similarity).
FT NP_BIND 264 286 PAPS (By similarity).
FT ACT_SITE 115 115 Proton acceptor (By similarity).
FT VARIANT 5 5 D -> E (in dbSNP:1402467).
FT /FTID=VAR 025404.
SQ SEQUENCE 302 AA; 35520 MW; 905E1820A5E222F3 CRC64;
Query Match 29.3%; Score 449; DB 1; Length 302;
Best Local Similarity 36.2%; Pred. No. 4.7e-29;
Matches 96; Conservative 52; Mismatches 99; Indels 18; Gaps 3;
QY 36 EIANFPVPSPDVWIVTPKSGTSLLOEVVYVLSQADPEIGLWNIDBQLPVLVYQPP- 94
DB 36 DKIMFQAKPDDLSTYPAKGTWTQEIIVELIQNEGVDEKSKRAPHQRPFFLEMKIPS 95
QY 95 ---GLDIIKELTSPRLIKSHLPYFLPSDLHNGSKVIYMARNPKDLVYVYQFHSRLT 151
DB 96 LQSGLEQAHMPSRLIKTHLPFLLPSSLLEKCKIIIVARNPKDNVSYTHFORMNKA 155
QY 152 MSYRGTQFQFCRRFNMNDKLGYSWEHFQEFWEHRMDSNVLFKYEDMHRDLVTVVQOLA 211
DB 156 LPAPGTWEYFETFLAGVCWSWEHVKGWWEAKDKHRLILFYEDMKKNPKHEIQKLA 215
QY 212 RFLGVSCDKAQLEALTEHCHQLV---DQCNAEALP-----VGRGVRGLWKDIF 257
DB 216 EFIGKLLDKVLDKIVHTSYFDVMQKNPMANYSSIPAEIMDSHSIPFMRKGAVGDKKHF 275
QY 258 TVSMNEKFDLVYKQKMGKCDLTFDF 282
DB 276 TVAQNERFEDYKKKMTDRLTFHP 300
RESULT 37
Q95MF8_PIG PRELIMINARY; PRT; 295 AA.
ID Q95MF8_PIG
AC Q95MF8;
DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT 01-DEC-2001, sequence version 1.
DT 07-FEB-2006, entry version 14.
DE Estrogen sulfotransferase.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
OC Sus.
OX NCBI_TaxID=9823;
RN [1]_
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22345564; PubMed=12457956; DOI=10.1016/S0739-7240(02)00172-8;
RA Kim J.G., Vallet J.L., Rohrer G.A., Christenson R.K.;
RT "Characterization of porcine estrogen sulfotransferase.";
RL Domest. Anim. Endocrinol. 23:493-506(2002).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonCommercial-ShareAlike
CC License.
DR EMBL; AF389855; AAK72967.1; -; mRNA.
DR HSSP; P49888; 1HY3.
DR SMR; Q95MF8; 4-293.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase 1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Transferase.
SQ SEQUENCE 295 AA; 34940 MW; 7677C28A91B38167 CRC64;
Query Match 29.2%; Score 447; DB 2; Length 295;
Best Local Similarity 34.3%; Pred. No. 6.7e-29;
Matches 96; Conservative 56; Mismatches 106; Indels 22; Gaps 4;
QY 16 ESKYFEF---HGVRLLPFPFCRGMKEIANPVRPSDVWIVTPKSGTSLLOEVVYVLSQ 71
DB 5 KSAVLDYFGRITGILLYKPFIEYWNVDVETPEARDDLVIVTPKSGTTWSEIVVMYTE 64
QY 72 ADPBEIGLWNIDBQLPVL---EYQPGCLDIIKELTSPRLIKSHLPYFLPSDLHNGDSK 127
DB 65 GDVEKCKEDTTFNRPYLECRTENVMNGVKQLQKQASPRIVKSHLPPELLEPVSWFKNCK 124
QY 128 VIYNARNPKDLVYVYQFHSRLTMSYRGTQFCRRFMNDKLGYSWFEHVFQEPWEHRM 187
DB 125 IIVCRNAKVVVSYIYFFFLMTANPDGSGQDFVEKFMDEVPYGVSKYTKSWWEKRT 184
QY 188 DSNVLFKYEDMHRDLVTVVQQLARFLGVSCDKAQLEALTEHCH-----QLV 234
DB 185 NPQVLFIYEDMKENIRKEVNRLEIFLGRKASDELVDKIHKHTSFQEMKNNPSTNYTLP 244
QY 235 DQCNAEALP-VGRGVRGLWKDIFTVSMNEKFDLVYKQKM 273
DB 245 DEVMNQKVSAPMRKGIAGDKWKNYFTVALNEKFDIHYEQQM 284
RESULT 38
ST1B1_RAT STANDARD; PRT; 299 AA.
ID ST1B1_RAT
AC P52847;
DT 01-OCT-1996, integrated into UniProtKB/Swiss-Prot.
DT 15-DEC-1998, sequence version 2.
DT 07-FEB-2006, entry version 42.
DE Sulfotransferase family cytosolic 1B member 1 (EC 2.8.2.-)
DE (Sulfotransferase 1B) (DOPA/tyrosine sulfotransferase).
GN Name=St1b1; Synonyms=St1b1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]_
RP NUCLEOTIDE SEQUENCE [MRNA], AND PROTEIN SEQUENCE OF 108-117; 126-134
RP AND 260-266.
RC STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=96107201; PubMed=8530477; DOI=10.1074/jbc.270.51.30470;
RA Sakakibara Y., Takami Y., Zwieb C., Nakayama T., Suiko M.,
RA Nakajima H., Liu M.-C.;
RT "Purification, characterization, and molecular cloning of a novel rat
RT liver Dopa/tyrosine sulfotransferase.";
RL J. Biol. Chem. 270:30470-30478(1995).
RN [2]_
RP NUCLEOTIDE SEQUENCE [MRNA].
RX MEDLINE=94306556; PubMed=8033246; DOI=10.1016/0009-2797(94)90057-4;
RA Yamazoe Y., Nagata K., Ozawa S., Kato R.;
RT "Structural similarity and diversity of sulfotransferases.";
RL Chem. Biol. Interact. 92:107-117(1994).
RN [3]_
RP NUCLEOTIDE SEQUENCE [MRNA].
RC TISSUE=Liver;
RX MEDLINE=98104061; PubMed=9443824;
RA Fujita K., Nagata K., Ozawa S., Sasano H., Yamazoe Y.;
RT "Molecular cloning and characterization of rat ST1B1 and human ST1B2
RT cDNAs, encoding thyroid hormone sulfotransferases.";
```

J. Biochem. 122:1052-1061(1997).
 (4)
 FUNCTION.
 RN PubMed=12773305; DOI=10.1152/ajpendo.00046.2003;
 RP Kaeter M.H., Kaptein E., Roest T.J., van Dijk C.H., Tibboel D.,
 RA Meinel W., Glatz H., Coughtrie M.W., Visser T.J.;
 RT "Characterization of rat iodothyronine sulfoxidation";
 RT Am. J. Physiol. 285:E592-E598(2003).
 CC -1- FUNCTION: Catalyzes the sulfate conjugation of many hormones,
 CC neurotransmitters, drugs and xenobiotic compounds. Sulfonation
 CC increases the water solubility of most compounds, and therefore
 CC their renal excretion, but it can also result in bioactivation to
 CC form active metabolites. Sulfonation of DOPA, tyrosine isomers and
 CC thyroid hormones such as 3,3',5-triiodothyronine and 3,3'-
 CC diiodothyronine. May play a role in the limitation of the
 CC production of L-DOPA and L-m-tyrosine and also in facilitating
 CC their excretion.
 CC -1- SUBUNIT: Monomer.
 CC -1- SUBCELLULAR LOCATION: Cytoplasm.
 CC -1- PTM: The N-terminus is blocked.
 CC -1- SIMILARITY: Belongs to the sulfotransferase family.
 CC
 CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC
 CC ENBL; U38419; AAC52387.1; -; mRNA.
 DR EMBL; D89375; BAA34546.1; -; mRNA.
 DR PIR; JC5884; JC5884.
 DR HSP; P50224; 1CJM.
 DR SMR; P52847; 5-296.
 DR Ensembl; ENSRGOG00000001967; Rattus norvegicus.
 DR RGD; 708534; Sult1b1.
 DR GO; GO:0005829; C:cytosol; TAS.
 DR GO; GO:0004062; F:aryl sulfotransferase activity; IDA.
 DR InterPro; IPR000863; Sulfotransferase.
 DR Pfam; PF00685; Sulfotransferase 1; 1.
 DR ProDom; PD001218; Sulfotransferase; 1.
 DR Direct protein sequencing; Lipid metabolism; Steroid metabolism;
 KW Transferrase.
 KW CHAIN 1 299 Sulfotransferase family cytosolic 1B
 member 1.
 FT FT
 FT NP_BIND 48 53 /FTID-PRO 0000085163.
 FT NP_BIND 131 139 PAPS (By similarity).
 FT NP_BIND 194 230 PAPS (By similarity).
 FT NP_BIND 109 109 PAPS (By similarity).
 FT ACT_SITE 259 259 Proton acceptor (By similarity).
 FT BINDING 259 259 PAPS (By similarity).
 FT CONFLICT 68 68 E -> G (in Ref. 1).
 FT SEQUENCE 299 AA; 34835 MW; FDECEC304EE788A8 CRC64;
 Query Match 29.2%; Score 446.5; DB 1; Length 299;
 Best Local Similarity 36.0%; Pred. No. 7.5e-29;
 Matches 95; Conservative 52; Mismatches 98; Indels 19; Gaps 3;
 Qy 36 EEIANFVPSDVWIVTYPKSGTSLLOEVVYVLSQADPDEIGLNMIDQLVLEYPQP- 94
 Db 29 EKIEEFSRQPCDIVITYPKSGTSLSEIVDMVNDGNVEKCKRDVITSKVPMLQNVPG 88
 Qy 95 -----GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVITYMARNPKDLVSVYQFHRSLR 150
 Db 89 ARRSGLVLEKTPSPRIIKTHLPIDLLPKSFWDNCKMTYLARNKGVAVSYTHFDLMNN 148
 Qy 151 TMSYRGTQFQFCRRFNMNDKLGVSWEFHEQWFEHWRMDSNVLFLKYEDMHRDLVTMVQL 210
 Db 149 IQPLPGTWEEYLEKFLAGNVAYGSWFDVKSWEKRGHPIFLFYEDDLKQPKKEIKKI 208
 Qy 211 ARFLVSCDKAQAEALTEHCHQLV-----DOCCNAEALP-----VGRGRVLGWDI 256
 Db 209 ANFLDKTLDHETLERIVHTTSFVEMKDNPLVNYTHLPTSDMHSKSPFMRKGVGDWKNY 268
 Qy 257 FTVSMNEKFDLVYKQKGGKDLTF 280
 Db 269 FTWTQSEKFDALYKKGSLGTTLEF 292

J. Biochem. 122:1052-1061(1997).
 (4)
 FUNCTION.
 RN PubMed=12773305; DOI=10.1152/ajpendo.00046.2003;
 RP Kaestner M.H., Kaptein E., Roest T.J., van Dijk C.H., Tibboel D.,
 RA Meinel W., Glatz H., Coughtrie M.W., Visser T.J.;
 RT "Characterization of rat iodothyronine sulfoxidation";
 CC Am. J. Physiol. 285:E592-E598(2003).
 CC -1- FUNCTION: Catalyzes the sulfate conjugation of many hormones,
 CC neurotransmitters, drugs and xenobiotic compounds. Sulfonation
 CC increases the water solubility of most compounds, and therefore
 CC their renal excretion, but it can also result in bioactivation to
 CC form active metabolites. Sulfonation of DOPA, tyrosine isomers and
 CC thyroid hormones such as 3,3',5-triiodothyronine and 3,3'-
 CC diiodothyronine. May play a role in the limitation of the
 CC production of L-DOPA and L-m-tyrosine and also in facilitating
 CC their excretion.
 CC -1- SUBUNIT: Monomer.
 CC -1- SUBCELLULAR LOCATION: Cytoplasm.
 CC -1- PTM: The N-terminus is blocked.
 CC -1- SIMILARITY: Belongs to the sulfotransferase family.
 CC
 CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC
 CC ENBL; U38419; AAC52387.1; -; mRNA.
 CC EMBL; D89375; BAA34546.1; -; mRNA.
 CC PIR; JC5884; JC5884.
 CC HSP; P50224; 1CJM.
 CC SMR; P52847; 5-296.
 CC Ensembl; ENSRGOG00000001967; Rattus norvegicus.
 CC RGD; 708534; Sult1b1.
 CC GO; GO:0005829; C:cytosol; TAS.
 CC GO; GO:0004062; F:aryl sulfotransferase activity; IDA.
 CC InterPro; IPR000863; Sulfotransferase.
 CC Pfam; PF00685; Sulfotransferase 1; 1.
 CC ProDom; PD001218; Sulfotransferase; 1.
 CC Direct protein sequencing; Lipid metabolism; Steroid metabolism;
 CC Transferrase.
 CC CHAIN 1 299 Sulfotransferase family cytosolic 1B
 member 1.
 FT FT
 FT NP_BIND 48 53 /FTID-PRO 0000085163.
 FT NP_BIND 131 139 PAPS (By similarity).
 FT NP_BIND 194 230 PAPS (By similarity).
 FT NP_BIND 109 109 PAPS (By similarity).
 FT ACT_SITE 259 259 Proton acceptor (By similarity).
 FT BINDING 68 68 PAPS (By similarity).
 FT CONFLICT 68 68 E -> G (in Ref. 1).
 FT SEQUENCE 299 AA; 34835 MW; FDECEC304EE788A8 CRC64;
 Query Match 29.2%; Score 446.5; DB 1; Length 299;
 Best Local Similarity 36.0%; Pred. No. 7.5e-29;
 Matches 95; Conservative 52; Mismatches 98; Indels 19; Gaps 3;
 Qy 36 EEIANFPVPSDVWIVTPKSGTSLLOEVVYVLSQADPDEIGLNMIDQLVLEYPQP- 94
 Db 29 EKIEEFSRQPCDIVTPKSGTSLSEIVDMVNDGNVEKCKRDVITSKVPMLQNVPG 88
 Qy 95 -----GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVITYMARNPKDLVSVYQFHRSLR 150
 Db 89 ARRSGLVKKTPSPRIIKTHLPIDLLPKSFWDNCKMTYLARNKGDVAVSYTHFDLMNN 148
 Qy 151 TMSYRTQFQFCRRFNMNDKLGVSWEFHEQWFEHWRMDSNVLFLKYEDMHRDLVTMVQQL 210
 Db 149 IQPLPGTWEEYLEKFLAGNVAYGSWFDVKSWEKRGHPIFLFYEDDLKQPKKEIKKI 208
 Qy 211 ARFLVSCDKAQAEALTEHCHQLV---DOCCNAEALP-----VGRGRVLGKDI 256
 Db 209 ANFLDKTLDHETLERIVHHTSFVEMKNDPLVNVYTHLPTSEIMHDSKSPFMRKGVGDWKNY 268
 Qy 257 FTVSMNEKFDLVYKQKGGKDLTF 280
 Db 269 FTWTQSEKFDALYKKGSLGTTLEF 292

FT	ACT SITE	108	108	Proton acceptor (By similarity).
FT	MOD_RES	216	216	Phosphoserine (by PKA) (Potential).
FT	MOD_RES	228	228	Phosphoserine (by PKA) (Potential).
FT	CONFLICT	117	117	S -> E (in Ref. 2).
FT	CONFLICT	171	171	S -> E (in Ref. 2).
FT	CONFLICT	248	248	M -> Q (in Ref. 2).
SEQ	SEQUENCE	295 AA;	34640 MW; 69147C73146913FD CRC64;	

Query Match	29.2%;	Score 446;	DB 1;	Length 295;
Best Local Similarity	34.2%;	Pred. No. 8.2e-28;		
Matches 100;	Conservative	57;	Mismatches 111;	Indels 24;
Gaps	5			

Qy	10	STFGPESEKFFEHGVRLLPFCFCGKMEIAN---	FPVRPSDVMVTVTVPKSGTSLLOEVVY	66
Db	3	SSPSPSDYEGKLGGI---	PWYKFFISQFINVEEFARPDDLVITVPKSGTTLWSEIC	59
Qy	67	LVSQGADPDPEIGLNMNIDEQFLPVL----	EYQPGLDIIKELTSPRLIKSHLPYRFLPDSLH	122
Db	60	MIYNGNDVEKCKEDVIFNRVPYLECSTEHVMKGVKQLNEMASPRIVKSHLPVKLLPVSWF	119	
Qy	123	NGDSKVIYMARNPKDLVSVSYQFHRSURTSWYRGTFQFCRCRRFNMNDKLGYSWFEHVOEF	182	
Db	120	EKNCKIYIISRNAKDVVSVYFFILMVTAPDPDSFQDFVEKFMDSGEVPGSWFEHTKSW	179	
Qy	183	WEHRMDSNVLFXYEDMHRDLVTMVEQLARFLGVSCDKAQLALETCH-----	231	
Db	180	WEKSNPQVQLFLPYEDMKNIRKVMKLLBFLGRKASDELVDKIKITKTSFOEMKKNPSTN	239	
Qy	232	--OLVQCCNAEALP-VGRGRVGLWKDIFTVMNKEFDLVYKQKMGKCDLTF	280	
Db	240	YTTLPDEVNMQKVSPPMRKGDVGDKWKHFTVALNEKEDMHYEQQMKSGTLKF	291	

RESULT 41	ST1B1 HUMAN	STANDARD;	PRT;	296 AA.
ID	ST1B1 HUMAN	STANDARD;	PRT;	296 AA.
AC	Q43704; O15497; Q96F11; Q9UK34;			
DT	04-JAN-2005,	integrated into UniProtKB/Swiss-Prot.		
DT	07-FEB-2006,	entry version 31.		
DE	Sulfotransferase family cytosolic 1B member 1 (EC 2.8.2.-)			
DE	(Sulfotransferase 1B2) (Thyroid hormone sulfotransferase).			
GN	Nome-SULT1B1; Synonyms=ST1B2, SULT1B2;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	TISSUE=Liver;			
RC	TISSUE=Liver;			
RX	MEDLINE=98104061; PubMed=9443824;			
RA	Fujita K., Nagata K., Ozawa S., Sasano H., Yamazoe Y.;			
RT	"Molecular cloning and characterization of rat ST1B1 and human ST1B2			
RT	cDNAs, encoding thyroid hormone sulfotransferases.";			
RL	J. Biochem. 122:1052-1061(1997).			
RN	[2]			
RP	TISSUE=Liver;			
RC	TISSUE=Liver;			
RX	MEDLINE=98130699; PubMed=9463486;			
RA	Wang J., Falany J.L., Falany C.N.;			
RT	"Expression and characterization of a novel thyroid hormone-sulfating			
RT	form of cytosolic sulfotransferase from human liver.";			
RL	Mol. Pharmacol. 53:274-282(1998).			
RN	[3]			
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].			
RC	TISSUE=Bone marrow;			
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;			
RA	Straubeberg R.L., Reinsgold E.A., Grouse L.H., Derge J.G.;			
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Altschul R.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RN	NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].			
RC	TISSUE=Bone marrow;			
RX	MEDLINE=98130699; PubMed=9463486;			
RA	Wang J., Falany J.L., Falany C.N.;			
RT	"Expression and characterization of a novel thyroid hormone-sulfating			
RT	form of cytosolic sulfotransferase from human liver.";			
RL	Mol. Pharmacol. 53:274-282(1998).			
RN	[3]			

Db 16 HGYPMTCAFASNWKIEQHSRPPDDIVATYPKSGTTWVSEIIMDLNDGDIKCKRGFI 75

QY 83 DEOLPVLVEYPOGL-----DIKELTSPLIKSHLPYRFLPSDLHNGDSKVIYMARPKD 137

Db 76 TEKVPMLEMTLPGLRTSGIEQLEKNPSRIVKTHLPDILLPKPSFWENNCKMIYLARNAXD 135

QY 138 LVVSYVOFHSRLTMSYRGTFQFCRRFMDNDKLGYSWFEHVOEFWEHRMDSNVLFKYE 197

Db 136 VVSYYTHFLDMNLQFPFGTWBEYLEKFLTKGVAGSWFTHVKNWKKKEHPILFLYYE 195

QY 198 DMHRDLVTWVEQLARFLGSCDKAQLEALTEHCHQLV---DOCCNAEALP----- 244

Db 196 DMKENPKEIKIIRFLKLNLDNDEILDRIIHTSFVEMKDNPLVNVYTHLPPTVMDHKSXP 255

QY 245 -VGRGRVGLWKOIFTVSMNEKEDLVYKQKMGKCDLTF 280

Db 256 FMRKGTAGDWKQNYFTVAQNEKFDALVETEMSKTALQF 292

RESULT 42

STIBI_CANFA

ID STIBI_CANFA STANDARD; PRT; 296 AA.

AC Q95JD5;

DT 04-JAN-2005, integrated into UniProtKB/Swiss-Prot.

DT 01-DEC-2001, sequence version 1.

DT 07-FEB-2006, entry version 22.

DE Sulfotransferase family cytosolic 1B member 1 (EC 2.8.2.-)

DE (Sulfotransferase 1B) (cSULT1B1).

GN Name=SULT1B1; Synonyms=ST1B2;

OS Canis familiaris (Dog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;

OC Canis.

OX NCBI_TaxID=9615;

UN NUCLEOTIDE SEQUENCE [MRNA], FUNCTION, AND TISSUE SPECIFICITY.

RP TISSUE=liver;

RC MEDLINE=21261348; PubMed=11368519; DOI=10.1006/abbi.2001.2373;

RA Tsou C., Falany C.N., Morgenstern R., Swedmark S.;

RT "Molecular cloning, expression, and characterization of a canine

RT sulfotransferase that is a human ST1B2 ortholog.;"

RL Arch. Biochem. Biophys. 390:87-92(2001).

CC -!- FUNCTION: Catalyzes the sulfate conjugation of many hormones,

CC neurotransmitters, drugs and xenobiotic compounds. Sulfonation

CC increases the water solubility of most compounds, and therefore

CC their renal excretion, but it can also result in bioactivation to

CC form active metabolites. Sulfates thyroid hormones including 3,3'-

CC diiodothyronine, 3,5,3'-triiodothyronine and 3,5'-diiodothyronine

CC (lesser degree) and phenols such as 1-naphthol and p-nitrophenol.

CC -!- SUBUNIT: Monomer (By similarity).

CC -!- SUBCELLULAR LOCATION: Cytoplasm (By similarity).

CC -!- TISSUE SPECIFICITY: Expressed highly in the colon, kidney and

CC small intestine of male and female dogs. Highly expressed in the

CC jejunum and ileum of the male dog than the female dog, which

CC displayed more expression in duodenum (at protein level).

CC -!- SIMILARITY: Belongs to the sulfotransferase family.

CC

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>

CC Distributed under the Creative Commons Attribution-NoDerivs License

CC

CC EMBL; AV004332; AAP86583.1; -; mRNA.

DR SMR; Q95JD5; 4-296.

DR Ensembl; ENSGAFG000000002871; Canis familiaris.

DR InterPro; IPR000863; Sulfotransferase.

DR Pfam; PF00685; Sulfotransferase 1; 1

DR ProDom; PD001218; Sulfotransferase; 1.

KW Lipid metabolism; Steroid metabolism; Transferase.

FT CHAIN 1 296 Sulfotransferase family cytosolic 1B member 1.

FT FT

FT NP_BIND 48 53 /FTID=PRO_0000085160.

FT NP_BIND 131 139 PAPS (By similarity).

FT NP_BIND 194 230 PAPS (By similarity).

FT FT

FT ACT_SITE 109 109 Proton acceptor (By similarity).

FT BINDING 259 259 PAPS (By similarity).

SQ SEQUENCE 296 AA; 34869 MW; 3D54CF003BF09365 CRC64;

Query Match 29.0%; Score 443; DB 1; Length 296;

Best Local Similarity 34.8%; Pred. No. 1.5e-28;

Matches 101; Conservative 51; Mismatches 118; Indels 20; Gaps 5;

QY 11 TPGEPSKYFB-FHGVRLPPFCRGKQKEBIANPPVRPSDVWIVTPYKSGTSLLOEVVYLV 69

Db 3 SPKDLRKNLKMIGHYPIIYTFANNWENIEQHSRPDDIIITATYPKSGTTWVSEIVDMVL 62

QY 70 QGADPEIGLNMNIDQLVLEYPQGL-----DIKELTSPLIKSHLPYRFLPSDLHNG 124

Db 63 NNGDVEKCKRDFITVKVPMLENAVPLGRTSGIEQLEKNPSRLVTHLPALLPKSFWE 122

QY 125 DSKVIYMARNPDKLVVSYQFHSRLTMSYRGTFQFCRRFMDNDKLGYSWFEHVOEFWE 184

Db 123 NCKMIYLARNAKDAVSYHFDLNMNLEPAPCPWEYLERFWTGNVAYGSWENHVKSWMK 182

QY 185 HRMDSNVLFKYEDMHRDLVTWVEQLARFLGSCDKAQLEALTEHC--HQLVDQ--CCNAE 241

Db 183 KKEHPILFLYEDMKENPKREVQKIARFLKLNLDNDEVLDKIIHTSFEMMKDNLVNYT 242

QY 242 ALP-----VGRGRVGLWKOIFTVSMNEKEDLVYKQKMGKCDLTF 280

Db 243 HLPSTVMDHKSXSFMRKGIAGDWKQNYFTVAQNEKFDVYKEMSGTTLQF 292

RESULT 43

STIBI_RABIT

ID STIBI_RABIT STANDARD; PRT; 296 AA.

AC Q46503;

DT 01-DEC-2000, integrated into UniProtKB/Swiss-Prot.

DT 01-JUN-1998, sequence version 1.

DT 07-FEB-2006, entry version 33.

DE Sulfotransferase 1C1 (EC 2.8.2.-) (rabsULT1C2).

GN Name=SULT1C1;

OS Oryctolagus cuniculus (Rabbit).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Glires; Lagomorpha; Leporidae;

OC Oryctolagus.

OX NCBI_TaxID=9986;

UN NUCLEOTIDE SEQUENCE [MRNA], AND CHARACTERIZATION.

RP TISSUE=Stomach;

RC MEDLINE=99410886; PubMed=10481272; DOI=10.1016/S1357-2725(99)00038-2;

RA Hehonah N., Zhu X., Brix L., Bolton-Grob R., Barnett A., Windmill K.,

RA McManus M.;

RT "Molecular cloning, expression, localisation and functional

RT characterisation of a rabbit SULT1C2 sulfotransferase.;"

RL Int. J. Biochem. Cell Biol. 31:869-882(1999).

CC -!- FUNCTION: Catalyzes the sulfate conjugation of many drugs,

CC xenobiotic compounds, hormones, and neurotransmitters. Shows

CC activity towards p-nitrophenol. Probably the major stomach

CC sulfotransferase.

CC -!- SUBCELLULAR LOCATION: Cytoplasm.

CC -!- TISSUE SPECIFICITY: Found in gastrointestinal tract tissues, liver

CC and kidney.

CC -!- SIMILARITY: Belongs to the sulfotransferase family.

CC

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>

CC Distributed under the Creative Commons Attribution-NoDerivs License

CC

CC EMBL; AF026304; AAC00410.1; -; mRNA.

DR HSSP; P50224; 1CJM.

DR SMR; Q46503; 12-296.

DR InterPro; IPR000863; Sulfotransferase.

DR Pfam; PF00685; Sulfotransferase 1; 1.

DR ProDom; PD001218; Sulfotransferase; 1.

KW Transferase.

FT CHAIN 1 296 Sulfotransferase 1C1.

FT /FTID=PRO_0000085134.

FT FT

RESN.T 45

RESULT 43
 Q6GP49_XENLA
 ID Q6GP49_XENLA PRELIMINARY; PRT: 297 AA.
 AC Q6GP49;
 DT 19-JUL-2004, integrated into UniProtKB/TrEMBL.
 DT 19-JUL-2004, sequence version 1.

OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidas;
OC Xenopodinae; Xenopus; Xenopus.
NCBI_TaxID=8355;
[1]
RN RT TISSUE=Splice;
RC
RX MEDLINE=22388257; Pubmed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonardo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uadin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Haile S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,
RA Fanev J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Groomed J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;
RT "generation and initial analysis of more than 15,000 full-length human
RL RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Spleen;
RA Klein S., Gerhard D.S.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.

```

CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BC073295.1; -; mRNA.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase_1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
SQ SEQUENCE 297 AA; 34860 MW; 88D37337F60C6603 CRC64;

Query Match      28.8%; Score 440; DB 2; Length 297;
Best Local Similarity 32.9%; Pred. No. 2.6e-28;
Matches 97; Conservative 53; Mismatches 115; Indels 30; Gaps 7;

QY 13 GBFEKVFPEF-----HGVRLPFFCR---GKMBEIANFPVRSDVWIVTPYKSGTSLLOQV 64
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2 GAQQLPYRFQLLPVHGI---PFMKPIADNWDRIETFOAKPDDLLIATYPKAGTTWMQBI 58
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 65 VYLVSGAGPDEIGLMNIDEQLPVLEY--POP---GLDIIKELTSRLIKSHLPYRFLPS 119
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 59 VDSIMNAGDLKKVGRAPTHVRFPFLICNPSPVPCGVDVLEETQSPRIIKTHLPYELIPK 118
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 120 DLHNGDSKVIYARNPKDLVSVYQPHRSRLRTMSYRGTFQECRRFMNDKLGYSWFHVV 179
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 119 SFWEHECKVIYARNAKDNVSYYPFDLMNKTQPHPTWEEYVGKLGKGNVPWGGFHHV 178
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 180 QEFWEHRMDSNVLFLKYEDMHRDLVTMVQLARFLGVSCDKAQLALTEHC---QLVDQ 236
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 179 IGWKKARAKHQILYFVEDMDKEDPKREIRKVRFLGKLSDELLEKICQHTSFKANKENP 238
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 237 CCNAEALP-----VGRGRVGLMKDIFTVSNMKEFDLVYKQKMGKCDLTF 280
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 239 MAYSAMPDVLQDSISPFMRKGEVSDWKVHFTVQQNMFDVDAEYQKRMGTDLKF 293
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 46
Q3TOY3 BOVIN PRELIMINARY; PRT; 296 AA.
AC Q3TOY3;
DT 11-OCT-2005, integrated into UniProtKB/TrEMBL.
DT 11-OCT-2005, sequence version 1.
DE 07-MAR-2006, entry version 4.
DE Hypothetical protein.
OS Bos taurus (Bovinae).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Crosbred x Angus; TISSUE=ileum;
RA Moore S.; Alexander L.; Brownstein M.; Guan L.; Lobo S.; Meng Y.;
RA Tanguchi M.; Wang Z.; Yu J.; Prange C.; Schreiber K.; Shenmen C.;
RA Wagner L.; Balu M.; Babazuk S.; Barber S.; Babakaiff R.; Beldand J.;
RA Chun E.; Del Rio L.; Gibson S.; Hanson R.; Kirkpatrick R.; Liu J.;
RA Matsuo C.; Mayo M.; Santos R.R.; Stott J.; Tsai M.; Wong D.;
RA Siddiqui A.; Holt R.; Jones S.J.; Marra M.A.;
RL Submitted (AUG-2005) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BC102208; AA102209.1; -; mRNA.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase_1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Hypothetical protein.
SQ SEQUENCE 296 AA; 34674 MW; 8A44653780078275 CRC64;

Query Match      28.5%; Score 436.5; DB 2; Length 296;
Best Local Similarity 33.9%; Pred. No. 5.1e-28;

```

```

Matches 94; Conservative 53; Mismatches 111; Indels 19; Gaps 4;

QY 23 HGVRLPFFCRGKMBEIANFPVRSDVWIVTPYKSGTSLLOQVYLVSGAGPDEIGLMNI 82
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 16 HGCPITYAFANNWEKIEQFSRPPDDIMIVTPYKSGTTSWISVDVLDHGDVCKCKRDI 75
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 83 DEQLPVLEYPOP-----GLDIIKELTSRLIKSHLPYRFLPSDLHNGDSKVIYARNPKD 137
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 76 TAKVEMLEALPGLRTSGLEKPNPSRVVKTLPDLIPKSFENNCKIIYLARNAD 135
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 138 LVSVYQPHRSRLRTMSYRGTFQECRRFMNDKLGYSWFHVEFWEHRMDSNVLFLKYE 197
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 136 VAVSYFHDLMNNLOPLPTGWGEVLEKELTGNVAGSFWNVKSWKKEGHPIFLFYE 195
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 198 DMHRDLVTMVQLARFLGVSCDKAQLALTEHC---HQLVDQ--CCNAEALP----- 244
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 196 DMKNPKQEKIKKVVRFLEKLNLDDELDKIYHTSFEMMKDNPVNYTHLPSEVMDHSS 255
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 245 -VGRGRVGLMKDIFTVSNMKEFDLVYKQKMGKCDLTF 280
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 256 FMRKGIAGDWKNYFTVAQNEKFDIYKEMSETELQF 292
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 47
STIE2 RAT
AC P52845;
DT 01-OCT-1996, integrated into UniProtKB/Swiss-Prot.
DT 01-OCT-1996, sequence version 1.
DT 07-FEB-2006, entry version 38.
DE Estrogen sulfotransferase, isoform 2 (EC 2.8.2.4) (EST-2)
DE (Sulfotransferase, estrogen-prefering) (Estrone sulfotransferase).
GN Name=Ste2; Synonyms=Ste;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA].
RC STRAIN=Fischer 344; TISSUE=Liver;
RX MEDLINE=96305357; PubMed=8688469; DOI=10.1016/0167-4781(96)00065-6;
RA Rikke B.A.; Roy A.K.;
RT "Structural relationships among members of the mammalian
RT sulfotransferase gene family.";
RL Biochim. Biophys. Acta 1307:331-338(1996).
CC -!- FUNCTION: Sulfation of estrone and estradiol. May control the
CC level of the estrogen receptor by sulfurylating free estradiol (By
CC similarity).
CC -!- CATALYTIC ACTIVITY: 3'-phosphoadenylyl sulfate + estrone =
CC adenosine 3',5'-bisphosphate + estrone 3-sulfate.
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasm (By similarity).
CC -!- SIMILARITY: Belongs to the sulfotransferase family.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; U50205; AA807681.1; -; mRNA.
DR HSSP; P49891; 1B06.
DR SMR; P52845; 7-294.
DR Ensembl; ENSRNOG00000001957; Rattus norvegicus.
DR RGD; 3776; Ste.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase_1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Lipid-binding; Steroid-binding; Transferase.
FT CHAIN 1 295 Estrogen sulfotransferase, isoform 2.
   /FTid=PRO_0000085156.
FT NP_BIND 48 53 PAPS (By similarity).
FT NP_BIND 130 138 PAPS (By similarity).
FT NP_BIND 193 229 PAPS (By similarity).
FT NP_BIND 257 259 PAPS (By similarity).

```


RT complex: evidence for catalytic role of Ser137 in the sulfonyl
 RT transfer reaction.";
 RL J. Biol. Chem. 277:17928-17932(2002).
 RN [9]
 RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) IN COMPLEX WITH ADENOSINE
 RP 3',5'-BISPHOSPHATE (PAP) AND A HYDROXYLATED POLYCHLORINATED BIPHENYL
 RP (OH-PCB).
 RX PubMed=12782487;
 RA Shevtsov S., Petrotchenko E.V., Pedersen L.C., Negishi M.;
 RT "Crystallographic analysis of a hydroxylated polychlorinated biphenyl
 RT (OH-PCB) bound to the catalytic estrogen binding site of human
 RT estrogen sulfotransferase.";
 RL Environ. Health Perspect. 111:884-888(2003).
 CC -|- FUNCTION: May control the level of the estrogen receptor by
 CC sulfonylating free estradiol. Maximally sulfates beta-estradiol
 CC and estrone at concentrations of 20 nM. Also sulfates
 CC dihydroepiandrosterone, pregnenolone, ethinylestradiol, equalenin,
 CC diethylstilbestrol and 1-naphthol, at significantly higher
 CC concentrations; however, cortisol, testosterone and dopamine are
 CC not sulfated.
 CC -|- CATALYTIC ACTIVITY: 3'-phosphoadenyl sulfate + estrone =
 CC adenosine 3',5'-bisphosphate + estrone 3-sulfate.
 CC -|- SUBUNIT: Homodimer.
 CC -|- SUBCELLULAR LOCATION: Cytoplasm.
 CC -|- TISSUE SPECIFICITY: Liver, intestine and at lower level in the
 CC kidney.
 CC -|- SIMILARITY: Belongs to the sulfotransferase family.
 CC
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC
 CC
 DR EMBL: U08098; AA82125.1; -; mRNA.
 DR EMBL: U20521; AAC50286.1; -; Genomic_DNA.
 DR EMBL: U20515; AAC50286.1; JOINED; Genomic_DNA.
 DR EMBL: U20516; AAC50286.1; JOINED; Genomic_DNA.
 DR EMBL: U20517; AAC50286.1; JOINED; Genomic_DNA.
 DR EMBL: U20518; AAC50286.1; JOINED; Genomic_DNA.
 DR EMBL: U20519; AAC50286.1; JOINED; Genomic_DNA.
 DR EMBL: U20520; AAC50286.1; JOINED; Genomic_DNA.
 DR EMBL: S77383; AA334601.1; -; mRNA.
 DR EMBL: Y11195; CAA72079.1; -; mRNA.
 DR EMBL: AY436634; AAQ97179.1; -; Genomic_DNA.
 DR EMBL: BC027956; AAH27956.1; -; mRNA.
 DR EMBL: U55764; AA951658.1; -; mRNA.
 DR F01R; J02229; J02229.
 DR PDB; 1G3M; X-ray; A/B=1-294.
 DR PDB; 1HY3; X-ray; A/B=1-294.
 DR IntAct; P49888; -.
 DR Ensembl; ENSG00000109193; Homo sapiens.
 DR H-InvDB; HIX0022597; -.
 DR HGNC; HGNC:11377; SULT1E1.
 DR MIM; 600043; Gene.
 DR GO; GO:0004304; F:estrogen sulfotransferase activity; TAS.
 DR GO; GO:0008202; P:steroid metabolism; TAS.
 DR InterPro; IPR000863; Sulfotransferase.
 DR Pfam; PF00685; Sulfotransferase_1; 1.
 DR ProDom; PD001218; Sulfotransferase; 1.
 DR 3D-structure; Lipid-binding; Polymorphism; Steroid-binding;
 KW Transferase.
 CHAIN 1 294 Estrogen sulfotransferase.
 /FTID=PRO_0000085153.
 FT NP_BIND 47 52 PAPS.
 FT NP_BIND 129 137 PAPS.
 FT NP_BIND 192 228 PAPS.
 FT NP_BIND 256 258 PAPS.
 FT ACT_SITE 107 107 Proton acceptor (By similarity).
 FT VARIANT 22 22 D -> Y (in dbSNP:11569705).
 FT /FTID=VAR_018907.
 FT MUTAGEN 137 137 S->A: Decreased gradually the catalytic
 FT activity.
 FT MUTAGEN 137 137 S->C: Decreased gradually the catalytic
 FT activity.
 FT MUTAGEN 269 269 V->E: Does not prevent the formation of

FT CONFLICT 154 154 homodimer.
 FT 10 F -> L (in Ref. 6).
 FT HELIX 4 14
 FT STRAND 11 14
 FT TURN 15 16
 FT STRAND 17 20
 FT HELIX 21 25
 FT TURN 26 26
 FT HELIX 27 31
 FT TURN 32 32
 FT TURN 37 38
 FT STRAND 40 45
 FT TURN 46 47
 FT STRAND 48 49
 FT HELIX 50 61
 FT TURN 62 64
 FT HELIX 66 69
 FT STRAND 70 72
 FT HELIX 74 77
 FT STRAND 78 78
 FT STRAND 80 80
 FT TURN 81 82
 FT STRAND 84 85
 FT TURN 86 88
 FT STRAND 89 90
 FT HELIX 91 96
 FT TURN 97 97
 FT STRAND 100 101
 FT STRAND 103 106
 FT HELIX 110 112
 FT STRAND 113 113
 FT HELIX 115 119
 FT TURN 120 121
 FT STRAND 123 128
 FT HELIX 131 144
 FT STRAND 145 145
 FT TURN 146 147
 Query Match 28.4%; Score 434; DB 1; Length 294;
 Best Local Similarity 33.3%; Pred. No. 8.2e-28;
 Matches 95; Conservative 56; Mismatches 116; Indels 18; Gaps 3;
 QY 14 EFESKYPEFHGVRLLPPFCRGMKEBIANPPVRPSPDVWIVTPKSGTSLIQEYVYLVSQAD 73
 Db 6 DYKEFEVHGILMYKDFVKYWDNVEAFQARPDDLVIATYPKSGTTWVSEIVYMYKEGD 65
 QY 74 PDEIGLMNIDQLPVLEYQPQ-----GLDIIKELTSPRIKSHLPYRFLPDLHNGDSKVI 129
 Db 66 VEKCKEDVIFNRIPFLPCRKENLMGVKQLDEMSPRIVKTHLPPELLPASFEWKDCKII 125
 QY 130 YMAENPKDLVSYQYQFHRSLTMSYRGTFQFCRREMNDKLGYSWFEHVQSFWEHRMDS 189
 Db 126 YLCRNADVAVSFYFFFLMVAHNPSPGPFPEFVKPMQGVPGVSGWKVKSMMWKGKSP 185
 QY 190 NVLFKLYEDMRDLVTWVEQLARFLGVSCDKAQLAEALTEHCH-----QLVDQ 236
 Db 186 RVLFELFYEDLAKEDIRKVKILHFLERKPSSELDVRIIHTSFQEMKNPNSTNYTLTDE 245
 QY 237 CCNAEALP-VGRGRVGLMKDIFTVSMNEKFDLVYKQMGKCDLTF 280
 Db 246 IMNQKLSPPMRKGITGDWKNHFTVALNEKFDKHYEQMKESTLKF 290
 RESULT 51
 Q53X91 HUMAN PRELIMINARY; PRT; 294 AA.
 ID Q53X91 HUMAN
 AC Q53X91;
 DT 24-MAY-2005, integrated into UniProtKB/TrEMBL.
 DT 24-MAY-2005, sequence version 1.
 DT 07-FEB-2006, entry version 3.
 DE SULT1E1 protein (fragment).
 GN Name=SULT1E1;
 OS Homo sapiens (Human).

OC Muroidea; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RA [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Methods Enzymol. 303:19-44(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX PubMed16141072; DOI=10.1126/science.1112014;
RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N., Oyama R., Ravasi T., Lenhard B., Wells C., Kodaira S., Yoshida H., Tanabe Y., Bickel S., Brenner S.E., Batalov S., Forrest A.R., Zavolan M., Davis M.J., Wilming L.G., Aldinis V., Allen J.E., Ambesi-Impiombato A., Apweiler R., Aturalien R.N., Bailei T.L., Banal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M., Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R., Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G., Di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G., Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M., Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E., Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N., Hill D., Hummel L., Iacono M., Ikeo K., Iwama A., Ishikawa T., Jakt M., Kanapin A., Katoh M., Kawasawa Y., Keiso J., Kitamura H., Kitano H., Kollas G., Krishnan S.P., Kruger A., Kummerfeld S.K., Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J., Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L., Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K., Motagui-Tabar S., Mulder N., Nakano N., Nakaguchi H., Ng P., Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O., Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavese G., Pesole G., Petrovsky N., Piazza S., Reed J.F., Ring B.Z., Ringwald M., Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C., Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sheng Y., Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B., Sperling S., Stupka E., Sugura K., Sultana R., Takenaka Y., Taki K., Tammoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A., Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K., Yamashiki H., Zabarovsky E., Zhai S., Zimmer A., Hide W., Bult C., Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J., Wahlesed C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y., Fukuda K., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T., Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N., Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Ninomiya N., Nishio T., Okada M., Plessy C., Shibata K., Shiraki T., Suzuki S., Tagami M., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J., Hayashizaki Y.;
RT "The transcriptional landscape of the mammalian genome.";
RL Science 309:1559-1563(2005).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX PubMed16141073; DOI=10.1126/science.1112009;
RG RIKEN Genome Exploration Research Group, and Genome Science Group (Genome Network Core Team) and the FANTOM Consortium;
RT "Antisense Transcription in the Mammalian Transcriptome.";
RL Science 309:1564-1566(2005).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S., Nikaide I., Oeato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H., Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T., Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J., Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W., Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S., Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S., Gaasterland T., Gariboldi M., Gissi C., Godik A., Gough J., Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D., Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L., Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A., Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H., Nagashima T., Numata K., Okido T., Pavan W.J., Perte G., Pesole G., Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S., Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M., Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K., Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M., Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C., Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L., Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N., Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K., Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S., Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kigawa I., Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A., Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J., Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S., Aizawa K., Iwama M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I., Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J., Schriml L.M., Stauber F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M., Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P., Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L., Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S., Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M., Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [7]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=20530913; PubMed=11078861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P., Konno H., Akiyama J., Nishi K., Kitanai T., Tashiro H., Itoh M., Suni N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A., Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K., Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M., Yoneda Y., Ishikawa T., Orawa K., Tanaka T., Matsura S., Kawai J., Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [8]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H., Arakawa T., Bono H., Carninci P., Fukuda S., Furuno M., Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,

```
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Nunazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AK015750; BAB29956.1; -; mRNA.
DR HSSP; P49891; 1B06.
DR SMR; Q9D566; 7-294.
DR Ensemble; ENSMUSG00000029272; Mus musculus.
DR MGI; MGI:98431; Sult1el.
DR GO; GO:0008210; P:estrogen metabolism; IMP.
DR GO; GO:0007565; P:pregnancy; IMP.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase 1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Transferase.
SQ SEQUENCE 295 AA; 35576 MW; 5526B3FC362BFB0E CRC64;

Query Match 28.2%; Score 432; DB 2; Length 295;
Best Local Similarity 33.0%; Pred. No. 1.2e-27;
Matches 94; Conservative 57; Mismatches 116; Indels 18; Gaps 3;

QY 14 EFESKYFEHGVRLPPFCRGKMEETANFPVRSDWIVTPYKSGTSLQEVVYLVSQAG 73
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
7 EYEVFGFGRGLVMDKRTFKYWEDEMFLARPDVIAATPKSGTWTSEVYMYIKEGD 66

QY 74 PDEIGLNMNIDEQLPVLEYPQP---GLDILIKELTSPRIKSHLPYRFLPSDLHNGDSKVI 129
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
67 VEKKEDAIENRIPYLECNEEDLINGIQKKEKESPRIVKTHLPVQLVPLFWRYCKV 126

QY 130 YMARNPKDLVSYVYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHFVQEFWEHRMDS 189
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
127 YLCRNAKDVAVSYYFFLLMITSYPNPKSFSEVPEKFMQGVPGYSWYDHVRAWEKSKNS 186

QY 190 NVFLFKYEDMRDLVTWVEQLARFLGVSCDCAQLEALTEHCH-----QLVDQ 236
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
187 RVLFMFYEDMKDIRREVYKLIETFLERKPSAELVDRIVQHTSFQEMKNPNSTNYTMPEE 246

QY 237 CCNAEALP-VGRGRVGLMKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
247 MNQKVSFPMRKGIIIGDWKNHFPPEALRERFDEHYKQMKDCTVKF 291

RESULT 54
Q5HZV7_XENTR PRELIMINARY; PRT; 287 AA.
ID Q5HZV7_XENTR
AC Q5HZV7;
DT 15-FEB-2005, integrated into UniProtKB/TrEMBL.
DT 15-FEB-2005, sequence version 1.
DT 07-FEB-2006, entry version 11.
DE Sulfotransferase family, cytosolic, 1A, phenol-preferring, member 1.
GN Name=sult1al;prov;
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus; Silurana.
OX NCBI_TaxID=8364;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=whole body;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
```

```
RL Dev. Dyn. 225:384-391 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole body;
RX MEDLINE=22389257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole body;
RA Klein S., Gerhardt D.S.;
RL Submitted (JAN-2005) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BC088870; AAH88870.1; -; mRNA.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR000963; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase 1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Transferase.
SQ SEQUENCE 287 AA; 33894 MW; 89CFB2DC7991479B CRC64;

Query Match 28.2%; Score 431.5; DB 2; Length 287;
Best Local Similarity 32.6%; Pred. No. 1.3e-27;
Matches 94; Conservative 58; Mismatches 115; Indels 21; Gaps 4;

QY 12 PGFESKYFEHGVRLPPFCRGKMEETANFPVRSDWIVTPYKSGTSLQEVVYLVSQG 71
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2 PGKLIQI-----VEGVSIAEDIASNQIQSQFOARPGDVLIATPKAGTTWQEIVDLIINE 57

QY 72 ADPDEIGLNMNIDEQLP---VLEYPQGLDIIKELTSPRIKSHLPYRFLPSDLHNGDSKV 128
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
58 GNEEICRRSPHTHERMPFVEVLHMMKPGPEEVNAMPSPRVLTHTLPVQLVPLFWRYCKV 117

QY 129 IYMARNPKDLVSYVYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHFVQEFWEHRMD 188
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
118 IYVARNPRDVTVSYYFDHTITFHPAPGSWEYLHFRFMKGDVGWGSWYDHVKGEWQKQD 177

QY 189 SNVLFLKYEDMRDLVTWVEQLARFLGVSCDCAQLEA---LTEHCHQLVDQCCNAEALP- 244
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
178 HNLIYLFYEDIKQNPPIHRIKRWRFDKDLSEEVLEKIVHLSSPFDHMKDNPMANFSAPPS 237

QY 245 -----VGRGRVGLMKDIFTVSMNEKFDLVYKQKMGKCDLTFDF 282
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
238 DVVDQSQYKFMKRGKVGDKWSHFTVQQNEMFEKYQQQMGHSAMKFRY 285

RESULT 55
ST1E1 MOUSE
ID ST1E1 MOUSE STANDARD; PRT; 295 AA.
AC P49891;
DT 01-OCT-1996, integrated into UniProtKB/Swiss-Prot.
```


DT 15-JUL-1999, sequence version 2.
 DT 07-PEB-2006, entry version 47.
 DE Estrogen sulfotransferase, testis isoform (EC 2.8.2.4)
 DE (Sulfotransferase, estrogen-prefering).
 GN Name=Sult1e1; Synonyms=Ste;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [mRNA].
 RC STRAIN=C57BL/KSJ-DB/DB; TISSUE=Testis;
 RX MEDLINE=95269690; PubMed=7750469; DOI=10.1210/en.136.6.2477;
 RA Song W.-C., Moore R., McLachlan J.A., Negishi M.;
 RT "Molecular characterization of a testis-specific estrogen
 sulfotransferase and aberrant liver expression in obese and
 diabetic C57BL/KSJ-db/db mice.";
 RL Endocrinology 136:2477-2484(1995).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS) IN COMPLEX WITH ADENOSINE
 3',5'-BISPHOSPHATE (PAP), AND SEQUENCE REVISION TO 113.
 RC STRAIN=C57BL/KSJ-DB/DB; TISSUE=Testis;
 RX MEDLINE=98025056; PubMed=9360604;
 RA Kakuta Y., Pedersen L.G., Carter C.W., Negishi M., Pedersen L.C.;
 RT "Crystal structure of estrogen sulphotransferase.";
 RL Nat. Struct. Biol. 4:904-908(1997).
 CC -!- FUNCTION: Sulfation of estradiol and estrone. May control the
 level of the estrogen receptor by sulfonylating free estradiol.
 CC -!- CATALYTIC ACTIVITY: 3'-phosphoadenylyl sulfate + estrone =
 adenosine 3',5'-bisphosphate + estrone 3-sulfate.
 CC -!- SUBUNIT: Homodimer [By similarity].
 CC -!- SUBCELLULAR LOCATION: Cytoplasm.
 CC -!- TISSUE SPECIFICITY: Testis and at very low level in the placenta.
 CC -!- MISCELLANEOUS: Abnormal high expression in liver in obese and
 diabetic C57BL/KSJ-DB/DB strain mice. Female > male. Normal
 level in liver.
 CC -!- SIMILARITY: Belongs to the sulfotransferase family.
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC -----
 DR EMBL; S78182; AAB34320.1; -; mRNA.
 DR PDB; 1AQU; X-ray; A/B=1-295.
 DR PDB; 1AQY; X-ray; A/B=1-295.
 DR PDB; 1BO6; X-ray; A/B=1-295.
 DR IntAct; P49891; -;
 DR Ensembl; ENSMUSG0000029272; Mus musculus.
 DR MGI; MGI:98431; Sult1e1.
 DR LinkHub; P49891; -;
 DR GO; GO:0008210; P:estrogen metabolism; IMP.
 DR GO; GO:0007565; P:pregnancy; IMP.
 DR InterPro; IPR000863; Sulfotransferase.
 DR Pfam; PF00685; Sulfotransferase 1; 1.
 DR ProDom; PD001218; Sulfotransferase; 1.
 KW 3D-structure; Lipid-binding; Steroid-binding; Transferase.
 FT CHAIN 1 295
 FT Estrogen sulfotransferase, testis
 FT Isoform.
 FT FT
 FT NP_BIND 48 53
 FT NP_BIND 130 138
 FT NP_BIND 193 229
 FT NP_BIND 257 259
 FT ACT_SITE 108 108
 FT CONFLICT 113 113
 FT HELIX 8 11
 FT STRAND 12 15
 FT TURN 16 17
 FT STRAND 18 21
 FT HELIX 22 24
 FT TURN 25 27
 FT STRAND 28 32
 FT TURN 33 33

FT TURN 38 39
 FT STRAND 41 45
 FT TURN 47 48
 FT STRAND 49 50
 FT HELIX 51 62
 FT TURN 63 64
 FT HELIX 75 78
 FT STRAND 79 79
 FT STRAND 81 81
 FT TURN 82 83
 FT STRAND 85 86
 FT TURN 87 89
 FT STRAND 90 91
 FT HELIX 92 97
 FT TURN 98 98
 FT STRAND 101 102
 FT STRAND 104 107
 FT HELIX 111 113
 FT STRAND 114 114
 FT HELIX 116 120
 FT TURN 121 122
 FT STRAND 124 129
 FT HELIX 132 145
 FT STRAND 146 146
 FT TURN 147 148
 FT STRAND 149 149
 FT STRAND 153 153
 FT HELIX 155 164
 FT TURN 165 165
 FT STRAND 167 167
 FT TURN 168 169
 FT HELIX 172 182
 FT TURN 183 184
 FT STRAND 185 185
 FT TURN 186 187
 FT STRAND 188 192
 FT HELIX 193 198
 FT TURN 200 210
 FT TURN 211 212
 FT HELIX 217 226
 FT TURN 227 227
 FT STRAND 228 228
 FT HELIX 229 234
 FT TURN 236 242
 FT STRAND 243 243
 FT TURN 245 247
 FT STRAND 248 248
 FT TURN 250 252
 FT STRAND 253 253
 FT STRAND 255 256
 FT TURN 261 261
 FT HELIX 263 266
 FT TURN 267 267
 FT HELIX 270 284
 FT TURN 285 286
 SQ SEQUENCE 295 AA; 35590 MW; 8585AB47952BFB1C CRC64;
 Query Match 28.2%; Score 431; DB 1; Length 295;
 Best Local Similarity 33.0%; Pred. NO. 1.5e-27;
 Matches 94; Conservative 57; Mismatches 116; Indels 18; Gaps 3;
 QY 14 EFESKYFEFHGVRLLPPFCRCGRKWEIAPPPVPSDVIIVTPKSGTSLAQEVVYLVSQAD 73
 Db 7 EYVEFGFRGLMDKRFKTKYTWEDVEMFLAPDDLVIAIYPSGTTIWEVVMYIKGD 66
 QY 74 PDEIGLNIQQLPVLEYPQP-GLDIKELTSPRIKSHLPYRFLPSDLHNGDSKV 129
 Db 67 VEKCKEDAIENRIPVLECRNEDLINGIKLAKESPRIVKTHLPKLLPASWEKNCMI 126
 QY 130 YNARNPKDLVSYVYQFHRSLRTMSYRGTFQFCRRFMDKLGYSWFEHQVFEHRMDS 189
 Db 127 YLCRNAKDVAVSYYFFLLIMITSYPNPKSFSEFVKFMQGPYPYGSWDHVKAWEKSKNS 186


```

QY 190 NVLFLKYEDMHRDLVTMTQARFLGVSCDKAQLALTEHCH-----QLVDQ 236
DB 187 RVLFMFYEDMKDIRREVVKLIEFLERKPSAELVDRIQHTSFQBMKNPSTNYTMMPEE 246

QY 237 CNAEALP-VGGRVGLWKDIFTVSNMKEFDLVYKQKMGKCDLTF 280
DB 247 MNQKVSFPMKGIIGDWKNHFPALRRFDEHYKQKMDCTVKF 291

RESULT 56
Q8JZX7 MOUSE
ID Q8JZX7_MOUSE PRELIMINARY; PRT; 295 AA.
AC Q8JZX7
DT 01-OCT-2002, integrated into UniProtKB/TrEMBL.
DT 01-OCT-2002, sequence version 1.
DT 07-FEB-2006, entry version 19.
DE Sulfotransferase family 1E, member 1.
GN Name=Sult1e1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
NUCLEOTIDE SEQUENCE.
RC TISSUE=Uterus;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M.J., Usdin T.B., Toshikiyuki S., Carninci P., Prange C.,
RA Brownstein M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Gay L.J., Hulyk S.W.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RN RP
NUCLEOTIDE SEQUENCE.
RC TISSUE=Uterus;
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
CC EMBL; BC034891; AAH34891.1; -; mRNA.
DR HSSP; P49891; 1A0Y.
DR SMR; Q8JZX7; 7-294.
DR Ensembl; ENSMUSG00000029272; Mus musculus.
DR MGI; MGI:98431; Sult1e1.
DR GO; GO:0008210; P:estrogen metabolism; IMP.
DR GO; GO:0007565; P:pregnancy; IMP.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
DR Transferrase.
SQ SEQUENCE 295 AA; 35590 MW; 5C2848FBD63E14F4 CRC64;

Query Match 28.2%; Score 431; DB 2; Length 295;
Best Local Similarity 33.0%; Pred. NO. 1.5e-27;
Matches 94; Conservative 57; Mismatches 116; Indels 18; Gaps 3;

```

```

QY 14 EFESKYFHFVRLPPFCRGKMEBIANPPVPSDWIIVTYPKSGTSLQEVVYLVSQAD 73
DB 7 EYEVFGFGRVGLWDRKFTKYWEDVEMFLARPDLLVIATYPKSGTTTSEVVYIYKEGD 66

QY 74 PDEIGLNMIDQLPVLEYPQP-GLDIIKELTSPRLIKSHLPVRFPLPSDLHNGDSKVI 129
DB 67 VEKCKEDAIENRIPLYECRNEDLINGIKQLAKESPRILKTHLPKLLPASFWKNCXMI 126

QY 130 YMARNPKDLVVSYYOFHRSRLTMSYRGTFQFCRRFMNDKLGYSWFFHVQEFWEHRMDS 189
DB 127 YLCRNAKDVAVSYVYFLIMITSYPNPKSFSFVEKFMQGVPGYSWYDHVKAWNEKSKNS 186

QY 190 NVLFLKYEDMHRDLVTMTQARFLGVSCDKAQLALTEHCH-----QLVDQ 236
DB 187 RVLFMFYEDMKDIRREVVKLIEFLERKPSAELVDRIQHTSFQBMKNPSTNYTMMPEE 246

QY 237 CNAEALP-VGGRVGLWKDIFTVSNMKEFDLVYKQKMGKCDLTF 280
DB 247 MNQKVSFPMKGIIGDWKNHFPALRRFDEHYKQKMDCTVKF 291

RESULT 57
Q99ND5 RAT
ID Q99ND5_RAT PRELIMINARY; PRT; 295 AA.
AC Q99ND5;
DT 01-JUN-2001, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2001, sequence version 1.
DT 07-FEB-2006, entry version 17.
DE Estrogen sulfotransferase.
GN Name=ste2;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
NUCLEOTIDE SEQUENCE.
RP Aspapova I.I., Yakovenko A.R., Shchelkunova T.A., Chernov B.K.,
RA Sverdlova P.S., Smirnov A.N., Rubtsov P.M.;
RT "Cloning and Preliminary Characterization of the Rat Estrogen
RL Sulfotransferase Gene 5'-Region.";
RL Mol. Biol. 33:447-453(1999).
[2]
RN RP
NUCLEOTIDE SEQUENCE.
RA Aspapova I.I., Yakovenko A.R., Morozov I.A., Smirnov A.N.,
RA Rubtsov P.M.;
RT "Cloning and Structural Characterization of Two Genes Encoding Rat
RT Estrogen Sulfotransferase.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
[3]
RN RP
NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshikiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Gay L.J., Hulyk S.W.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

```

```

RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Liver;
RC Director MGC Project;
RA Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
RL J. Biochem. 124:55-64(1998).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AJ306223; CAC27405.3; JOINED; Genomic DNA.
DR EMBL; AJ306224; CAC27405.3; JOINED; Genomic DNA.
DR EMBL; AJ306225; CAC27405.3; JOINED; Genomic DNA.
DR EMBL; AJ306226; CAC27405.3; JOINED; Genomic DNA.
DR EMBL; AJ306227; CAC27405.3; JOINED; Genomic DNA.
DR EMBL; AJ306228; CAC27405.3; JOINED; Genomic DNA.
DR EMBL; AJ306229; CAC27405.3; JOINED; Genomic DNA.
DR EMBL; BC088157; AAB88157.1; mRNA.
DR HSP; P49891; 1AQU.
DR SMR; Q99ND5; 7-294.
DR Ensembl; ENSRNOG0000001957; Rattus norvegicus.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Transferase.
SQ SEQUENCE 295 AA; 35362 MW; 6A470865BA8F5B0E CRC64;

Query Match 28.1%; Score 430; DB 2; Length 295;
Best Local Similarity 31.5%; Pred. No. 1.8e-27;
Matches 92; Conservative 62; Mismatches 106; Indels 32; Gaps 4;

QY 14 EFSEKYFEHGVRLPPFCRGKMEETANFPSPDVWITYPKSGTSLQEVVYLVSQAD 73
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 7 EYEVFGFHGVLMDKLTFTKYWEFTSFARDDLLVTPKSGTSLQEVVYLVSQAD 66
QY 74 PDEIGLMDIEOLPVLVYQPQ----GLDIIKELTSPRIKSLHPLRPLSPDLNGDSKVI 129
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 67 VEKKEDAIKFNPIFYECRNEDLNGIKQLKEKSPRIKTHPLPKLPAFWEKNCKII 126
QY 130 YNARPKDLVSYOYFHRSLRTSYRGTFQEFCCRFRMNDKLGYSWFEHVSQEFWEHRMDS 189
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 127 YLCRAKADVWSYVYFFLIISYPNPKSPFSEPVKEFMGQVPGYSGWYHVSWEKSKNS 186
QY 190 NVFLPKYEDMRDLVTWVEQLARFLVGSQCDKRAQLTEHCHQLV-----DQCEN 239
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 187 RVLFMFYEDMKEDIRREVVKLTIEFL-----ERDPLAELVDKIITHTSFQEMKNPNCTN 239
QY 240 AEALP-----VGRGVGLWKDIFTVSMNEKFDLVYKOKMGKCDLTF 280
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 240 YSMLPETWIDLKVSFPMKRGVGDWRNHPFPAIRERFHEHYQRHMKDCPVKF 291

RESULT 58
ST1B1 MOUSE STANDARD; PRT; 299 AA.
AC Q9QW67; Q9C301; Q9ZT0;
DT 04-JAN-2005, integrated into UniProtKB/Swiss-Prot.
DT 04-JAN-2005, sequence version 2.
DE 07-FEB-2006, entry version 29.
DE Sulfotransferase family cytosolic 1B member 1 (EC 2.8.2.-)
DE (Sulfotransferase 1B) (DOPA/tyrosine sulfotransferase).
GN Name=St1b1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), FUNCTION, TISSUE SPECIFICITY,
RP AND DEVELOPMENTAL STAGE.
RC STRAIN=C57BL/6 X CBA; TISSUE=Liver;
RX PubMed=9644246;
```

```

RA Saeki Y., Sakakibara Y., Araki Y., Yanagisawa K., Suiko M.,
RA Nakajima H., Liu M.-C.;
RT "Molecular cloning, expression, and characterization of a novel mouse
RT liver SUL1B1 sulfotransferase.";
RL J. Biochem. 124:55-64(1998).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORMS 1 AND 2).
RC STRAIN=C57BL/6J; TISSUE=Oviduct;
RX PubMed=16141072; DOI=10.1126/science.1112014;
RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
RA Oyama R., Ravasi T., Lenhard B., Wells K., Kodzius R., Shimokawa K.,
RA Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,
RA Davis M.J., Wilming L.G., Aidinis V., Allen J.E.,
RA Ambesi-Impiombato A., Apweiler R., Auraliysa R.N., Bailey T.L.,
RA Bansal M.P., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
RA di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G.,
RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,
RA Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,
RA Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
RA Hill D., Huminecki L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
RA Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,
RA Kitano H., Kollias G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
RA Matsuda H., Matsuzawa S., Mulder N., Nakano N., Nakautchi H., Ng P.,
RA Mottaqui-Tabar S., Mulder N., Nakano N., Nakauchi H., Ng P.,
RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavese G., Pesole G.,
RA Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
RA Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,
RA Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sheng Y.,
RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
RA Sperling S., Stupka E., Sugiuira K., Sultana R., Takenaka Y., Taki K.,
RA Tammoja K., Tan S.L., Tang S., Taylor M.J., Tegner J., Teichmann S.A.,
RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,
RA Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,
RA Wallemstedt C., Matick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
RA Kawashima T., Koike M., Kondo S., Konno H., Nakano K., Nimomiya N.,
RA Nishio T., Okada M., Plesky C., Shibata K., Shiraki T., Suzuki S.,
RA Tagami T., Waki K., Watanabe A., Okamura-Oho Y., Suzuki H., Kawai J.,
RA Hayashizaki Y.;
RT "The transcriptional landscape of the mammalian genome.";
RL Science 309:1559-1563(2005).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).
RC STRAIN=FVB/N; TISSUE=Colon;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.2426038999;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butcherfield V.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Catalyzes the sulfate conjugation of many hormones,
```

—
••
••
—
••
••
—
—
—

[illegible]

```
QY 6 AETPTGCEFSKYFEFHGVRLLPPFCRCGMEEIANFPVRPSDVMTVTYKSGTSLLOQV 65
Db 1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2 ACPNENMEPLKPLVPVIGIPLIKYFAETMQLQNTAHPDDVLITTPKSGTWMSEIM 61
QY 66 YLVSGADPDEIGLNMIDEQLFVLEYP-----QPGLDIIKELTSPRLKSHLPYRFLPSDL 121
Db 1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
62 DMIVGGKLDKGRAPVYARIPFLFSCPGVPPGLETLKETPAPRIKTHPLPLSLPQSL 121
QY 122 HNGDSKVYMARNPDLVSVYQFHRSLRTWSYRGTFCRCFRFNDKLGVGSWFHFVQE 181
Db 1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
122 LQDKIKVYVARNADVVVYFNFKMAKJHPDGTWESFLENFMGDKVSGVSWYQHVKE 181
QY 182 FWEHRMDSNVLFKYEDMHRDLVTWVEOLARFLGVSCDKAOLEALTEH 229
Db 1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
182 MWELRTHPVLYFYEDMKENPKREIKKILEFLGRSLPEETVDLIVHH 229

RESULT 60
ST1E3 RAT STANDARD; PRT; 295 AA.
AC F49889; Q9QWS0;
DT 01-OCT-1996, integrated into UniProtKB/Swiss-Prot.
DT 01-OCT-1996, sequence version 1.
DT 07-FEB-2006, entry version 41.
DE Estrogen sulfotransferase, isoform 3 (EC 2.8.2.4) (EST-3)
DE (Sulfotransferase, estrogen-prefering) (Estrone sulfotransferase).
GN Name=Ste;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP TISSUE=Liver;
RC NUCLEOTIDE SEQUENCE, AND PROTEIN SEQUENCE OF 258-265.
RX MEDLINE=92261615; PubMed=1374839; DOI=10.1210/me.6.4.589;
RA Demyan W.P., Song C.S., Kim D.S., Her S., Gallwitz W., Rao T.R.,
RA Slonczynska M., Chatterjee B., Roy A.K.;
RT "Estrogen sulfotransferase of the rat liver: complementary DNA cloning
RT and age- and sex-specific regulation of messenger RNA.";
RL Mol. Endocrinol. 6:589-597(1992).
RN [2]
RP TISSUE=Liver;
RC NUCLEOTIDE SEQUENCE.
RX MEDLINE=95161323; PubMed=7857871; DOI=10.1016/0960-0760(94)00147-E;
RA Falany J.L., Kraenkh V., Mikhneva G., Falany C.N.;
RT "Isolation and expression of an isoform of rat estrogen
RT sulfotransferase.";
RL J. Steroid Biochem. Mol. Biol. 52:35-44(1995).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=12173467;
RA Astapova I.I., Smirnov A.N., Rubtsov P.M.;
RT "PCR amplification and structural analysis of two paralogous rat
RT estrogen sulfotransferase genes.";
RL Mol. Biol. (Mosk.) 36:635-642(2002).
CC -!- FUNCTION: Sulfation of estrone and estradiol. May control the
CC level of the estrogen receptor by sulfonylating free estradiol.
CC -!- CATALYTIC ACTIVITY: 3'-phosphadenylyl sulfate + estrone =
CC adenosine 3',5'-bisphosphate + estrone 3-sulfate.
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasm.
CC -!- TISSUE SPECIFICITY: Liver of young mature males and uterus.
CC -!- DEVELOPMENTAL STAGE: Expressed only in the liver of young adult
CC animals (100 days old) and is absent in the prepubertal male (27
CC days old), senescent male (800 days old) and female liver.
CC -!- INDUCTION: Induced by androgens and suppressed by estrogens. The
CC expression is under the influence of pituitary growth hormone and
CC thyroid hormone. Is regulated by progesterone in the uterus.
CC -!- SIMILARITY: Belongs to the sulfotransferase family.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
```

```
CC EMBL; M86758; AAA41128.1; -; mRNA.
DR EMBL; S76489; AAB33441.1; -; mRNA.
DR EMBL; AJ131835; CAA10515.2; -; Genomic DNA.
DR EMBL; AJ298109; CAA10515.2; JOINED; Genomic DNA.
DR EMBL; AJ298110; CAA10515.2; JOINED; Genomic DNA.
DR EMBL; AJ298111; CAA10515.2; JOINED; Genomic DNA.
DR EMBL; AJ298112; CAA10515.2; JOINED; Genomic DNA.
DR EMBL; AJ298113; CAA10515.2; JOINED; Genomic DNA.
DR EMBL; AJ298114; CAA10515.2; JOINED; Genomic DNA.
DR PIR; A41930; A41930.
DR HSSP; P49891; 1AQU.
DR SMR; P49889; 7-294.
DR Ensembl; ENSRNOG0000001957; Rattus norvegicus.
DR RGD; 3776; Ste.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Direct protein sequencing; Lipid-binding; Steroid-binding;
KW Transferase.
FT CHAIN 1 295 Estrogen sulfotransferase, isoform 3.
FT NP_BIND 48 53 /FTID=PRO_0000085157.
FT NP_BIND 130 138 PAPS (By similarity).
FT NP_BIND 193 229 PAPS (By similarity).
FT NP_BIND 257 259 PAPS (By similarity).
FT ACT_SITE 108 108 Proton acceptor (By similarity).
FT CONFLICT 150 150 P -> Q (in Ref. 2).
FT CONFLICT 238 238 T -> I (in Ref. 2).
FT CONFLICT 295 295 L -> P (in Ref. 2).
SQ SEQUENCE 295 AA; 35416 MW; A77807A21DD2E7EB CRC64;

Query Match 28.0%; Score 429; DB 1; Length 295;
Best Local Similarity 32.3%; Pred. No. 2.2e-27;
Matches 92; Conservative 60; Mismatches 115; Indels 18; Gaps 3;

QY 14 EFESKYFEFHGVRLLPPFCRCGMEEIANFPVRPSDVMTVTYKSGTSLLOQVYLVSGAD 73
Db 7 EYDVFGDFHGLMDKRFYKWEVETFLARPDLILVTPKSGSTWSEIVDMYKGD 66
QY 74 PDEIGLNMIDEQLFVLEYPQP-----GLDIIKELTSPRLKSHLPYRFLPSDLHNGDSKVI 129
Db 67 VEKCKEDALFNRIPOLECRNEDLINGIKQLKEKESPRIVKTHLPKLLPASFWKCNCKII 126
QY 130 YMAENPDLVSVYQFHRSLRTWSYRGTFCRCFRFNDKLGVGSWFHFVQEFWHRMDS 189
Db 127 YLCRNADVVVSYYYFFFLIMKSYPNPKSFSFVEKFMEGQVPYGSWYDHVSWMEKSKNS 186
QY 190 NVLFLKYEDMHRDLVTWVEOLARFLGVSCDKAQLAEALTEHC---HQLVDQCCNAEALP-- 244
Db 187 RVLFMFVYEDMKEDIRREVVKLIEFLERDPSAELVDRIIQTSTFQEMKNNPCTNYSMLPET 246
QY 245 -----VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 247 MDLKVSPFMKRGIVGDKNHPPEALRERFEEHYQQQMKDCPVKF 291

RESULT 61
Q3T0S9 BOVIN
ID Q3T0S9_BOVIN PRELIMINARY; PRT; 295 AA.
AC Q3T0S9;
DT 11-OCT-2005, integrated into UniProtKB/TrEMBL.
DT 11-OCT-2005, sequence version 1.
DT 07-MAR-2006, entry version 4.
DE Hypothetical protein.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Crossbred x Angus; TISSUE=ileum;
```

[illegible]


```

RESULT 65
Q4FZP1_XENLA
ID Q4FZP1_XENLA PRELIMINARY; PRT; 302 AA.
AC Q4FZP1;
DT 30-AUG-2005, integrated into UniProtKB/TrEMBL.
DT 30-AUG-2005, sequence version 1.
DT 07-FEB-2006, entry version 2.
DE Hypothetical protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole;
RC TISSUE=Whole;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.,
RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny K.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalios D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole;
RA Klein S., Gethard D.S.;
RL Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BC099307; AAH99307.1; -; mRNA.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Hypothetical protein.
SQ SEQUENCE 302 AA; 35792 MW; AF91BD3FB8BF90C9 CRC64;

Query Match 27.9%; Score 427.5; DB 2; Length 302;
Best Local Similarity 31.6%; Pred. No. 3e-27;
Matches 89; Conservative 61; Mismatches 115; Indels 17; Gaps 3;

QY 18 KYFEPHGVRLPPFCRGKMEIANFPVPSPDVWIVTPKSGTSLLOEVVLYVLSQCADPDEI 77
ID Q6IM16_HUMAN PRELIMINARY; PRT; 304 AA.
DB Q6IM16;
QY 78 GLMNIDEQLP---VLEYQPQGLDIKELTSPRLKSHLPYRFLPSDLHNGDSKVIMARN 134

RESULT 67
Q6IM16_HUMAN
ID Q6IM16_HUMAN PRELIMINARY; PRT; 304 AA.
AC Q6IM16;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.

```


DT 05-JUL-2004, sequence version 1.
 DE 07-FEB-2006, entry version 10.
 GN SULT1C3 splice variant d.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX PubMed=14676822; DOI=10.1038/gt.tj.6500223;
 RA Freimuth R.R., Weipert M., Chute C.G., Wieben E.D., Weinshilboum R.M.;
 RT "Human cytosolic sulfotransferase database mining: identification of
 seven novel genes and pseudogenes."
 RL Pharmacogenomics J. 4:54-65(2004).
 CC -!- MISCELLANEOUS: The sequence shown here is derived from an
 EMBL/GenBank/DBJ third party annotation (TPA) entry.
 CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC -----
 DR EMBL; BK001432; DAA01771.1; -; Genomic DNA.
 DR HSP; P49891.1A0V.
 DR Ensembl; ENSG00000196228; Homo sapiens.
 DR GO; GO:0008146; F:sulfotransferase activity; IEA.
 DR InterPro; IPR000863; Sulfotransferase.
 DR Pfam; PF00685; Sulfotransferase_1; 1.
 DR ProDom; PD001218; Sulfotransferase; 1.
 SQ SEQUENCE 304 AA; 35889 MW; 5510C869578BB70C CRC64;

 Query Match 27.6%; Score 422; DB 2; Length 304;
 Best Local Similarity 32.3%; Pred. No. 8.7e-27;
 Matches 97; Conservative 59; Mismatches 124; Indels 20; Gaps 5;

 QY 1 MAESAETPTSGEPE-SKYEFHGVRLPPFCRCGMKEETANPPVPSPDVWIVTPKSGTS 59
 DB 1 MAKIEKAPTMEKKELFNEMVDGVTLLSKWEKVCNFOAKPDLILATYPSGTT 60

 QY 60 LQEVVYLVQSQADPEIGLNMIDQLPVLE--YP---OPGLDIIKEITSPRLIKSHLPY 114
 DB 61 WHHEILMDLNDGVEKCKRAQTLDRHAFLELKFPHKEKPDLEFVLEMSSPQILKTHLPS 120

 QY 115 RPLPSDLHNGDSKVLYMARNKDLVYVYQFHRSLRTWSYRGTFQFCRRRNMKLGYS 174
 DB 121 HLIPSPWKENCKIYVVARNPCKDLVSYHFRHMASFMPDPPQNLBEFYEKFNKGVVGS 180

 QY 175 WPEHVQEFWEHRMDSNVLPFLKYEDMHRDLVTWVEQLARFLGVSCDQAQLEALTEHCHQLV 234
 DB 181 WFDHVKGWAAKDMHRLVLFYEDIKKOPKRIEKLKFLKXDISSEILNKIIYHTSFDV 240

 QY 235 ---DQCNAEALPVG-----RGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
 DB 241 MKQNPMTNVTTLPTSIMDHSISPFMRKMGPGDKNYFTVAQNEEFKDYQKMGAGSTLTF 300

 RESULT 68
 ID Q6XZC1 BRARE PRELIMINARY; PRT; 304 AA.
 AC Q6XZC1
 DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
 DT 05-JUL-2004, sequence version 1.
 DE 07-FEB-2006, entry version 15.
 DE SULT1 sulfotransferase isoform 4.
 GN Name=sult1s4; ORFNames=zgc:76922;
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Liu M.-C., Liu C.-C., Sugahara T.;

RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Kidney;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards D., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield V.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Kidney;
 RG NIH MGC Project;
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
 CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC -----
 DR EMBL; AY196986; AAP55638.1; -; mRNA.
 DR EMBL; BC066584; AAH66584.1; -; mRNA.
 DR HSP; P49891.1A0V.
 DR Ensembl; ENSDARG0000003181; Danio rerio.
 DR ZFIN; ZDB-GENE-040426-2054; sult1s4.
 DR GO; GO:0008146; F:sulfotransferase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR InterPro; IPR000863; Sulfotransferase.
 DR Pfam; PF00685; Sulfotransferase_1; 1.
 DR ProDom; PD001218; Sulfotransferase; 1.
 KW Transferase.
 SQ SEQUENCE 304 AA; 35556 MW; 20BE69D9DC409873 CRC64;

 Query Match 27.5%; Score 421.5; DB 2; Length 304;
 Best Local Similarity 34.9%; Pred. No. 9.6e-27;
 Matches 98; Conservative 51; Mismatches 111; Indels 21; Gaps 6;

 QY 20 FEHGVRLPPFCRCGMKEETANPPVPSPDVWIVTPKSGTSLQEVVYLVQSQADPEIGL 79
 DB 21 FDFEGVLTREFTDNWENVKNFOARDDILATYPKAGTTWVSYLTLFVSGDENQTS- 79

 QY 80 MNIDEQLPVLE-----YQPGLDIIKEL-TSPRLIKSHLPVRLPLPSDLHNGDSKVYMA 132
 DB 80 QPIVORVPFLSCFQEFSTIGTEMADNLPTSPRLIKTHLPVLQVPKSFWEQNSRVVYA 139

 QY 133 RNPKDLVSVYQFHRSLRTWSYRGTFQFCRRRNMKLGYSWFEHVOEFWE-HRMSDNV 191
 DB 140 RNAKDNVSYFHFDRMNMVQDPDGDWSDYLDKFMQGVNFGVSWFDVSGWQKRSYPMN 199

 QY 192 LFLKYEDMHRDLVTWVEQLARFLGVSCDQAQLEALTEHCH-----QLVDQCNAEALP---- 244
 DB 200 LYMFEDSEDTGREVNRKLCFLGLSTSVQEKITKGVQVQFDAMKQNTLINHTVTPFLDC 259

 QY 245 -----VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
 DB 260 KISPFMRKGVGDWKSHTTVAQNERFDEVYKQKMGKNSGVTF 300

 RESULT 69

OC Muroidea; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
[1]
RP NUCLEOTIDE SEQUENCE [MRNA], AND PROTEIN SEQUENCE OF 10-20 AND 28-38.
RC STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=94043328; PubMed=8227031;
RA Nagata K., Ozawa S., Miyata M., Shimada M., Gong D.-W., Yamazoe Y.,
RA Kato R.;
RT "Isolation and expression of a cDNA encoding a male-specific rat
RT sulfotransferase that catalyzes activation of N-hydroxy-2-
RT acetylaminofluorene.";
RL J. Biol. Chem. 268:24720-24725(1993).
[2]
RN NUCLEOTIDE SEQUENCE [MRNA].
RP MEDLINE=9430556; PubMed=8033246; DOI=10.1016/0009-2797(94)90057-4;
RX Yamazoe Y., Nagata K., Ozawa S., Kato R.;
RA "Structural similarity and diversity of sulfotransferases.";
RL Chem. Biol. Interact. 92:107-117(1994).
CC -!- FUNCTION: Sulfation of phenols and bioactivation of N-
CC hydroxyarylamines. Is responsible for the formation of N-hydroxy-
CC 2-acetylaminofluorene, a reactive metabolite which exhibits
CC toxicity by binding to DNA, RNA and protein.
CC -!- SUBCELLULAR LOCATION: Cytoplasm (potential).
CC -!- TISSUE SPECIFICITY: Liver. Male >> Female.
CC -!- DEVELOPMENTAL STAGE: Male specific. Maximum at 9 weeks and
CC maintained in 9-month-old rats. Can be detected at low level in
CC females up to 9-week-old rats but then decreases to undetectable
CC level.
CC -!- INDUCTION: Induced by estrogens and suppressed by androgens.
CC -!- SIMILARITY: Belongs to the sulfotransferase family.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL: L22339; AAA42181.1; -; mRNA.
DR PIR: A49098; A49098.
DR HSSP: P50224; 1CJM.
DR Ensembl: ENSRNOG00000011269; Rattus norvegicus.
DR LinkHub: P50337; -.
DR GO: GO:0005829; Cytosol; NAS.
DR GO: GO:004062; F:aryl sulfotransferase activity; IDA.
DR InterPro: IPR000863; Sulfotransferase.
DR Pfam: PF00685; Sulfotransfer_1; 1.
DR ProDom: PD001218; Sulfotransferase; 1.
KW Direct protein sequencing; Transferase.
FT CHAIN 1 304 N-hydroxyarylamines sulfotransferase.
FT /FTID=PRO 0000085138.
FT NP_BIND 56 61 PAPS (By similarity).
FT NP_BIND 139 147 PAPS (By similarity).
FT NP_BIND 202 238 PAPS (By similarity).
FT NP_BIND 266 268 PAPS (By similarity).
FT ACT_SITE 117 117 Proton acceptor (By similarity).
SQ SEQUENCE 304 AA; 35764 MW; C772B2EA7BD74198 CRC64;
Query Match 27.4%; Score 419.5; DB 1; Length 304;
Best Local Similarity 32.6%; Pred. No. 1.4e-26;
Matches 91; Conservative 57; Mismatches 11; Indels 19; Gaps 4;
QY 21 EFHGVRLPPFCRGKMEIANFPVPSDVMIVTPYKSGTSLQEVVYLVVSQADPDEIGLM 80
DB 22 EVNGILMSKLSMDNDWKIWNFOAKPDDLIIATYAKAGTTWTQEIIVDMIQNDGVDQKQRA 81
QY 81 NIDQLPVLEY--POP---GLDIKELTSPRLIKSHLPYRFLPSDLNGLNDSKVIYMARNP 135
DB 82 NYDRHPPIETWLPSPNLNSGLDLANKMPSRPTLTKHLPVHMLPPSFWKENSKIIVARNA 141
QY 136 KDLVSVYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVEHWRMDSNVLFK 195
DB 142 KDLVSVYIFFRMNKMLDPDGTGLGVIIEQFAGKVLGWSWDVHVGWWDVQDQHRILYLF 201
QY 196 YEDMRDLVTMVQELARFLGVSCDKAQLEALTECHQLV---DQCNEALP-----244
DB 202 YEDNKEDPKRIKIAKLEKDISBEVLNKIIYHTSFDVMKENPMANVTTLTLPSSIMDSHI 261

QY 245 ---VGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTF 280
DB 262 SPFMRKMGPGDKNKFYFTVAQSEDFEDYRRKMAGSNITF 300
RESULT 72
Q5M8B5 RAT PRELIMINARY; PRT; 304 AA.
ID Q5M8B5 RAT
AC Q5M8B5;
DT 01-FEB-2005, integrated into UniProtKB/TrEMBL.
DT 01-FEB-2005, sequence version 1.
DT 07-FEB-2006, entry version 8.
DE Sultic1 protein.
GN Name=Sultic1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muroidea; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
[1]
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Liver;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uesdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RG NIH MGC Project;
RL Submitted (DSC-2004) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL: BC088125; AAH88125.1; -; mRNA.
DR GO: GO:0008146; F:sulfotransferase activity; IEA.
DR InterPro: IPR000863; Sulfotransferase.
DR Pfam: PF00685; Sulfotransfer_1; 1.
DR ProDom: PD001218; Sulfotransferase; 1.
SQ SEQUENCE 304 AA; 35774 MW; 961E7D24B9B8FOA CRC64;
Query Match 27.4%; Score 419.5; DB 2; Length 304;
Best Local Similarity 32.3%; Pred. No. 1.4e-26;
Matches 90; Conservative 57; Mismatches 113; Indels 19; Gaps 3;
QY 21 EFHGVRLPPFCRGKMEIANFPVPSDVMIVTPYKSGTSLQEVVYLVVSQADPDEIGLM 80
DB 22 EVNGILMSKLSMDNDWKIWNFOAKPDDLIIATYAKAGTTWTQEIIVDMIQNDGVDQKQRA 81
QY 81 NIDQLPVLEYPOP-----GLDIKELTSPRLIKSHLPYRFLPSDLNGLNDSKVIYMARNP 135
DB 82 NYDRHPPIETWLPSPNLNSGLDLANKMPSRPTLTKHLPVHMLPPSFWKENSKIIVARNA 141
QY 136 KDLVSVYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVEHWRMDSNVLFK 195

```

Db 142 KDLVSYYYFSRNKMLPDGTLGEYIEQFKAGKWLGSWYDHWKGMWDVKDQHRILYLF 201
QY 196 YEDMRDLVTMVEQLARFLGVSCDKAQLAEALTEHCHQLV---DQCNAEALP----- 244
Db 202 YEDMKEDPKREIKKAKFLKLEKIDSEVLNKLIIYHVSFDVWKENPMANYTTLSSIMDSHI 261
QY 245 ----VGRGRVGLWKDIFTVSMNKKFPLVYKQKMGKCDLTF 280
Db 262 SPFMKMGMDGWNFTVAQSEDFDDEYRRKMAGSNITF 300

RESULT 73
Q68P84_XENTR PRELIMINARY; PRT; 287 AA.
AC Q68P84;
DT 05-JUL-2004, integrated into UniProtKB/T-EMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 14.
DE Hypothetical protein MGC75696.
GN Name=MGC75696;
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus; Silurana.
OX NCBI_TaxID=8364;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Frange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RA Klein S., Gerhard D.S.; to the EMBL/GenBank/DBSJ databases.
RL Submitted (NOV-2003) to the EMBL/GenBank/DBSJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BC061263; AAH61263.1; -; mRNA.
DR HSSP; F49891; IAQU.
DR Ensemble; ENSXGTG00000012821; Xenopus tropicalis.
DR GO; GO:0008146; F:sulfofransferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransfer 1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Hypothetical protein.
SQ SEQUENCE 287 AA; 33145 MW; F5F382031752B4A8 CRC64;

Query Match 27.2%; Score 416.5; DB 2; Length 287;
Best Local Similarity 34.6%; Pred. No. 2.3e-26;
Matches 94; Conservative 48; Mismatches 111; Indels 19; Gaps 4;

```

```

QY 27 LPPFCRGKQVBIANFPVPRPSDVWIVTYPKSGTSLLOEVVYLVSQGADPDEIGLMNIDROL 86
Db 15 LGPFA-ANWENVKFEQARPDLLIATYPKSGTWMSEIVDQIVAVNSERCCKTAIYERV 73
QY 87 PVLEYPOP---GLDIIKELTSPRLIKSHLPVRELPSPDLHNGDSKVIYMARPKDLVSY 142
Db 74 PFLEYAVPDMPSGTQALDQRASPRLIKTHLPVELLPKSFWDNKVKVIIVARNKDVAVSY 133
QY 143 YQPHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEEHVFWEHRMDSNVFLFKYEDMHRD 202
Db 134 YHFYRMAIVHPPEPGTWDFLDSYINGKVCFGSWSAHVKGWQKAKEMDVLYLFYEDMLED 193
QY 203 LVTWVEQLARFLGVSCDKAQLAEAL---TEHCHQLVDQCNAEALP-----VGRG 248
Db 194 PTREIRKVKVFMGKDLPEETVEKIASQTSFKAMKQNELSNYSWVPSSVMDHSISPFRMG 253
QY 249 RVGLWKDIFTVSMNKKFPLVYKQKMGKCDLTF 280
Db 254 VGDWKNQFTVAQNEKFDDEYQREMSDGLSIF 285

RESULT 74
Q68EV4_XENLA PRELIMINARY; PRT; 276 AA.
AC Q68EV4;
DT 11-OCT-2004, integrated into UniProtKB/T-EMBL.
DT 07-FEB-2006, entry version 12.
DE MGC84291 protein.
GN Name=MGC84291;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Frange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain;
RA Klein S., Gerhard D.S.; to the EMBL/GenBank/DBSJ databases.
RL Submitted (AUG-2004) to the EMBL/GenBank/DBSJ databases.

```


Db 150 PEPGTWEELEKPMAGQSGPWYDHVKGWWEKREYRILYLFYEDMKENPKCEIQOILK 209

QY 213 FLGVSCDKAQLEALTEHCH-----QLVDQCCNABALPVGGRVGLWKDIFT 258

Db 210 YLEKDIPEELINKILYHSSFSVMKENPSANYITMKEEMDHSVPFMRKGISGDWKNQFT 269

QY 259 VSNKEKFDLVYKQKMGKCDLTF 280

Db 270 VAQYKFEEDYVKMKMEDSTLKF 291

RESULT 77

QSEAWO XENLA PRELIMINARY; PRT; 303 AA.

AC QSEAWO;

DT 15-MAR-2005, integrated into UniProtKB/TrEMBL.

DT 07-FEB-2006, entry version 1.

DE Hypothetical protein.

OS Xenopus laevis (African clawed frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;

OC Xenopodinae; Xenopus; Xenopus.

OX NCBI_TaxID=8355;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC TISSUE=Egg;

RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;

RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,

RA Richardson P.;

RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus

RT initiative.;"

RL Dev. Dyn. 225:384-391 (2002).

RN [2]

RP NUCLEOTIDE SEQUENCE.

RC TISSUE=Egg;

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,

RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,

RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences.;"

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

RN [3]

RP NUCLEOTIDE SEQUENCE.

RC TISSUE=Egg;

RA Klein S., Gerhard D.S.;

RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>

CC Distributed under the Creative Commons Attribution-NoDerivs License

CC -----

DR EMBL; BC090224; AAH90224.1; -; mRNA.

DR GO; GO:0008146; F:sulfotransferase activity; IEA.

DR InterPro; IPRO00863; Sulfotransferase.

DR Pfam; PF00685; Sulfotransfer 1; 1.

DR ProDom; PD001218; Sulfotransferase; 1.

KW Hypothetical protein.

SQ SEQUENCE 303 AA; 35647 MW; B940559678751DA1 CRC64;

Query Match 27.0%; Score 413; DB 2; Length 303;

Best Local Similarity 32.1%; Pred. No. 4.9e-26;

Matches 89; Conservative 55; Mismatches 115; Indels 18; Gaps 3;

QY 22 FHGVRLLPPFCRCMBEIANFPVRPSDVJWTVYKSGTSLQLQVVLVSGADPDEIGLWN 81

Db 23 FRGVPMAKMNVENWEPVEYFQARHDDVVLATYPKAGTTWVSEILDIMYNGDLEKQORD 82

QY 82 IDEQLPVLEYPOGL-----DIIKELTSPRLIKSHLPYRPLPSDLHNGDSKVLYMARNP 137

Db 83 IYNRVPMIEIRIPGMPGVDQLELLASPLRIKTHLPQLMPSEFWEKKCKVLYVARNAKD 142

QY 138 LVVSYVQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVQEFWEHRMDSNVLPFKYE 197

Db 143 VAVSYFFHFQMVKALPDGPDWKLADYMGVTSYGSWYDHVKGWWEKREKQYGLYLFYE 202

QY 198 DMHRDLVTWVEQLARFLGVSCDKAQLEALTEHCHQLV-----DOCCNABALP----- 244

Db 203 DLKEDPKREIKKILHYLEKELSDVELEKIVVHRTSFQVMSKNDWMANYKTIPNDILNOTNTA 262

QY 245 -VGRGRVGLWKDIFTVSNKEKFDLVYKQKMGKCDLTF 280

Db 263 FMKGEAGDWKNHFTVAQNEAFDNOYQBOMLGTSLHF 299

RESULT 78

Q80VR3_MOUSE PRELIMINARY; PRT; 304 AA.

AC Q80VR3;

DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.

DT 01-JUN-2003, sequence version 1.

DT 07-FEB-2006, entry version 18.

DE Sulfotransferase family, cytosolic, 1C, member 1.

GN Name=Sult1c1;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;

OC Muridea; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC TISSUE=Olfactory epithelium;

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,

RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences.;"

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

RN [2]

RP NUCLEOTIDE SEQUENCE.

RC TISSUE=Olfactory epithelium;

RA Director MGC Project;

RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>

CC Distributed under the Creative Commons Attribution-NoDerivs License

CC -----

DR	ENBL; BC045149; AAH45149.1; -; mRNA.
DR	HSSP; P50224; ICM.
DR	Ensembl; ENSMUSG00000023943; Mus musculus.
DR	MGI; MGI:102928; Sult1c1.
DR	GO; GO:0005737; C:cytoplasm; IDA.
DR	GO; GO:0004062; F:aryl sulfotransferase activity; IDA.
DR	GO; GO:0006790; P:sulfur metabolism; IDA.
DR	InterPro; IPR00863; Sulfotransferase.
DR	Pfam; PF00685; Sulfotransferase_1; 1.
DR	ProDom; PD001218; Sulfotransferase; 1.
KW	Transferase.
SQL	SEQUENCE 304 AA; 35797 MW; 9CF4A4B63713B977 CRC64;
Query Match	27.0%; Score 412.5; DB 2; Length 304;
Best Local Similarity	31.9%; Pred. No. 5.4e-26;
Matches	89; Conservative 56; Mismatches 115; Indels 19; Gaps 3
QY	21 EPHGVRLPPFCRGKVEIANFPVPSDVIVTYPKSGTSLQEVVYLVSQCADPDEIGLM 80
DB	22 EVNGILMSKMSSEMDKINWFQAKPDDLLIATYAKAGTWTQEIVMDQDGVQKQORA 81
QY	81 NIDQLPVLEYQPQ-----GLDIIKEITSPRLIKSHLPVRLPSDLHNGDSKVIYVARNP 135
DB	82 NTYDRHPIETWLPPLNSGLDLANKMSPRTLKTHLPVQMLPPSPFWKENSIIYVARNA 141
QY	136 KDLVVSYYYQFHRSLRTWSYRGTFQFCRRFMNDKLGYSWFHQQEFWEHRMDSNVLFK 195
DB	142 KCLVSYYYFSRNKMKLPDPCTLGEIYETFKAGKVLGWSYDVKHGWVDVKHRLVLF 201
QY	196 YEDMRDLVTWVEQLARPLGVSCDKAQLEALTEHCHQLV----DQCNAEALP----- 244
DB	202 YEDMKDKPKREIKKIVKPLEKDISEVLNKIHHSTFDVMKQNPMANYTTLPSSIMDHSI 261
QY	245 ---VGRGRVLGKWDIFTVSMNEKEDLVYKQKMGKCDLTF 280
DB	262 SPFMRKMGPGDKWNYFTVAQSEDFEDYRKKNAGSTITF 300
RESULT 79	
STIAL_BOVIN	STANDARD; PRT; 294 AA.
ID	AC P50227;
DT	01-OCT-1996, integrated into UniProtKB/Swiss-Prot.
DT	01-OCT-1996, sequence version 1.
DT	07-FEB-2006, entry version 43.
DE	Sulfotransferase 1a1 (BC 2.8.2.1) (Aryl sulfotransferase) (Phenol
DE	sulfotransferase) (Phenol-sulfating phenol sulfotransferase) (P-PST).
GN	Name=SULT1A1; Synonyms=STP;
OS	Bos taurus (Bovine).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC	Cecora; Bovidae; Bovinae; Bos.
OX	NCBI_TaxID=9913;
ON	[1]
RP	NUCLEOTIDE SEQUENCE, AND PARTIAL PROTEIN SEQUENCE.
RP	TISSUE=Tracheobronchial;
RX	MEDLINE=96003918; PubMed=7575456;
RA	Schauss S.J., Henry T., Palmatier R., Halvorson L., Dannenbring R.,
RA	Beckmann J.D.;
RT	"Characterization of bovine tracheobronchial phenol sulphotransferase
RT	cDNA and detection of mRNA regulation by cortisol.";
RL	Biochem. J. 311:209-217(1995).
RL	[2]
RP	NUCLEOTIDE SEQUENCE.
RX	PubMed=8890738; DOI=10.1016/0378-1119(96)00083-2;
RA	Henry T., Kliewer B., Palmatier R., Ulphani J.S., Beckmann J.D.;
RT	"Isolation and characterization of a bovine gene encoding phenol
RT	sulfotransferase.";
RL	Gene 174:221-224(1996).
RL	[3]
RP	NUCLEOTIDE SEQUENCE OF 24-57.
RA	Nonnenan D.J., Shibuya H., Johnson G.S.;
RL	Submitted (DEC-1994) to the ENBL/GenBank/DBJ databases.

MEDLINE=23289257; PubMed=12477932; DOI=10.1073/pnas.242603899;					
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh P., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Locoellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McQuellan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.C., Sodergren E.J., Lu X., Gibbs R.A., Patney J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield V.S.N., Krzywinski M.I., Skalska U., Smailus D.E., Schnorch A., Schein J.E., Jones S.J.M., Marra M.A.; "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."				
RT	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).				
RN	[2]				
RP	NUCLEOTIDE SEQUENCE.				
RC	TISSUE=Whole;				
RX	MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;				
RA	Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W., Richardson P.;				
RA	"Genetic and genomic tools for Xenopus research: The NIH Xenopus Initiative";				
RT	Dev. Dyn. 225:384-391 (2002).				
RL	[3]				
RN	NUCLEOTIDE SEQUENCE.				
RP	TISSUE=Whole;				
RC	Klein S., Strausberg R.;				
RA	Submitted (JUN-2003) to the ENBL/GenBank/DBJ databases.				
CC	-----				
CC	Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms				
CC	Distributed under the Creative Commons Attribution-NoDerivs License				
CC	-----				
DR	EMBL; BC053792; AAHS3792.1; -; mRNA.				
DR	HSSP; P49888; 1HY3.				
DR	GO; GO:0008146; F:sulfotransferase activity; IEA.				
DR	InterPro; IPR000863; Sulfotransferase.				
DR	Pfam; PF00685; Sulfotransferase 1.				
DR	ProDom; PD001218; Sulfotransferase: 1.				
SQ	SEQUENCE 294 AA; 34217 MW; 9DA265CAE40BE476 CRC64;				
 Query Match Best Local Similarity 31.2%; Score 412; DB 2; Length 294; Matches 91; Conservative 58; Mismatches 121; Indels 22; Gaps 4;					
QY	13 GEFE---SKYFEEFGHVLPPFCRGCMEEIANFPVPSPDVWIVTPKSGTSLLOEVVLYS 69				
Db	3 GELBWRKKDDVVGVPIIAAFNSNWERIKNFQARADDIVCTYPKSGTTWISDIVDL 62				
QY	70 QGADPDDEIGLMNIDQLPVLEYPOPLD-----I IKELTSPLRIKSHLPYRFLPSDLHG 124				
Db	63 SDGDTDKSRDAIHKKVPMLEFSAPGGVASSGLVLESVPSPRMKTHLTVSLLPKSFWEK 122				
QY	125 DSKVIYWARNPOLVSYYYQPHRSLRMTWSYRGTFQFCRRFRMNDKLGYSGWFHVQEFWE 184				
Db	123 KCKYVYVARNPDKVAVSFYHPDKMNQLHPPEGPMDKYLEKFMQGVKGVGPGMPHVRDWE 182				
QY	185 HRMDSNVLFPLYEDWHRDLVTMVQLARPLGVSCDKAQLEALTEHC-----OLVQCQNAE 241				
Db	183 LRKKQNMLFYFEDIEDPKREIRKVITSFKGLDPETIVEKICQHTSPKAMKENPLTNYS 242				
QY	242 ALP-----VGRGVLGWKDIFTVSMNEKFDLVYKMGKCDLTDFD 282				
Db	243 SVPSAVMDQSIPFMRKIAGRWNRHFTFAOESRFDEYVEGEVAATDLUSRF 294				
RESULT 81					

RESULT 81

[illegible]

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Muridae; Murinae; Mus.
 OK NCBI_taxid=10090;
 RN [1]
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Ovary and uterus;
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning.";
 RL Methods Enzymol. 303:19-44(1999).
 RN [2]
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Ovary and uterus;
 RX PubMed=16141072; DOI=10.1126/science.1112014;
 RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
 RA Oyama R., Ravasi B., Lenhard B., Wells C., Kodaira K., Tomoda K.,
 RA Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,
 RA Davis M.J., Wilming L.G., Aidinis V., Allen J.E.,
 RA Ambesi-Impombato A., Apweiler R., Aturaliya R.N., Bailey T.L.,
 RA Bansal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
 RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
 RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
 RA di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G.,
 RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gabibidi M.,
 RA Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,
 RA Guscinich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
 RA Hill D., Hummel M., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
 RA Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,
 RA Kitano H., Kollas G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
 RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
 RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
 RA Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
 RA Mottagui-Tabar S., Mulder N., Nakano N., Nakachi H., Ng P.,
 RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
 RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavese G., Pesole G.,
 RA Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
 RA Roost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,
 RA Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sheng Y.,
 RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
 RA Sperling S., Stupka E., Sugtara K., Sultana R., Takenaka Y., Taki K.,
 RA Tammoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
 RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,
 RA Yamani H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,
 RA Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
 RA Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
 RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
 RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
 RA Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Ninomiya N.,
 RA Nishio T., Okada M., Plessy C., Shibata K., Shiraki T., Suzuki S.,
 RA Tagami M., Waki K., Watanabe A., Okamura-Oho Y., Suzuki H., Kawai J.,
 RA Hayashizaki Y.;
 RT "The transcriptional landscape of the mammalian genome.";
 RL Science 309:1559-1563(2005).
 RN [3]
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Ovary and uterus;
 RX PubMed=16141073; DOI=10.1126/science.1111009;
 RG RIKEN Genome Exploration Research Group, and Genome Science Group
 RT (Genome Network Core Team) and the FANTOM Consortium;
 RT "Antisense Transcription in the Mammalian Transcriptome.";
 RL Science 309:1564-1566(2005).
 RN [4]
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Ovary and uterus;
 RX MEDLINE=23354683; PubMed=12466851; DOI=10.1038/nature01266;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nikaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,

GAasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Konagaya A., Kurochkin I.V., Lee Y., Lennard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Perte G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wyshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kigawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RT Bitney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RN [5]
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Ovary and uterus;
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Iwama M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Hashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 403:685-690(2001).
 RN [6]
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Ovary and uterus;
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630(2000).
 RN [7]
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Ovary and uterus;
 RX MEDLINE=2530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
 RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kaehiwagi K.,
 RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanabe M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771(2000).
 RN [8]
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Ovary and uterus;
 RA Arakawa T., Carninci P., Fukuda S., Hashizume W., Hayashida K.,

RA Hori F., Iida J., Imamura K., Imotani K., Itoh M., Kanagawa S.,
RA Kawai J., Kojima M., Konno H., Murata M., Nakamura M., Ninomiya N.,
RA Nishiyori H., Nomura K., Ohno M., Sakazume N., Sano H., Sasaki D.,
RA Shibata K., Shiraki T., Tagami M., Tagami Y., Waki K., Watahiki A.,
RA Muramatsu M., Hayashizaki Y.,
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AK133530; BAE21709.1; -; mRNA.
DR MGI; MGI:1926341; Sult1d1.
DR GO; GO:0004062; F-aryyl sulfotransferase activity; IDA.
DR GO; GO:0000103; P-sulfate assimilation; IDA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase 1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Transferase.
SQ SEQUENCE 295 AA; 35083 MW; A60A9AF60CC2736F CRC64;
Query Match 26.9%; Score 412; DB 2; Length 295;
Best Local Similarity 31.7%; Pred. No. 5.7e-26;
Matches 83; Conservative 63; Mismatches 98; Indels 18; Gaps 2;
QY 37 ETANFPVPSDWIYTPKSGTSLQEVVYLVSQADPDEIGLMDIDQLPVLVYPQGL 96
DB 30 QVESFEARPDILLISTYPKSGTWTWSEILDLYNNGDAEKCKRDALYKRVPFMELIIPGI 89
QY 97 ----DIKELTSPRIKSHLPVRLPDLHNGDSKVIYMARNPDLVSYVQFHRSLRTM 152
DB 90 TNGVEMLNMPSPRIVKTHLPVQLLPSSFVNKNDCKLIIVARNAKDVVSYFYQMAKIH 149
QY 153 SYRGTFQBFRCRRFMNDKLGYSWFHEVQEFWEHRMDSNVLFKYEDMHRDLVTWVEQLAR 212
DB 150 PEPGTWSEFLEKFMAGQVSFGPWYDHVKSWEKREYRILYLFYEDMKENPKCEIQKILK 209
QY 213 FLGVSCDKAQLAEALTEHCH-----QLVDQCCNAEALPVGRVGLMKDIFT 258
DB 210 FLEKDIPEILNKILYHSFSVMKGNPSANYTTMMKEEMDHVSFPMRKGISGDWKNQPT 269
QY 259 VSMNEKFDLVYKQKMGKCDLTF 280
DB 270 VAQYEKFEEDYVKRMEDSTLKF 291
RESULT 83
Q6NZD1_MOUSE PRELIMINARY; PRT; 295 AA.
AC Q6NZD1_MOUSE
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 13.
DE Sulfotransferase family 1D, member 1.
GN Name=Sult1d1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=C57BL/6J; TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner K.H., Shenmen C.M., Schuler G.D.,
RA Altchul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshikiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., McEwan P.J., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RA Director MGC Project;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BC066190; AAH66190.1; -; mRNA.
DR HSSP; P49891; IAQU.
DR Ensembl; ENSMUSG00000029273; Mus musculus.
DR MGI; MGI:1926341; Sult1d1.
DR GO; GO:0004062; F-aryyl sulfotransferase activity; IDA.
DR GO; GO:0000103; P-sulfate assimilation; IDA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase 1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Transferase.
SQ SEQUENCE 295 AA; 35011 MW; A569FAF60CC0716F CRC64;
Query Match 26.9%; Score 412; DB 2; Length 295;
Best Local Similarity 31.7%; Pred. No. 5.7e-26;
Matches 83; Conservative 63; Mismatches 98; Indels 18; Gaps 2;
QY 37 ETANFPVPSDWIYTPKSGTSLQEVVYLVSQADPDEIGLMDIDQLPVLVYPQGL 96
DB 30 QVESFEARPDILLISTYPKSGTWTWSEILDLYNNGDAEKCKRDALYKRVPFMELIIPGI 89
QY 97 ----DIKELTSPRIKSHLPVRLPDLHNGDSKVIYMARNPDLVSYVQFHRSLRTM 152
DB 90 TNGVEMLNMPSPRIVKTHLPVQLLPSSFVNKNDCKLIIVARNAKDVVSYFYQMAKIH 149
QY 153 SYRGTFQBFRCRRFMNDKLGYSWFHEVQEFWEHRMDSNVLFKYEDMHRDLVTWVEQLAR 212
DB 150 PEPGTWSEFLEKFMAGQVSFGPWYDHVKSWEKREYRILYLFYEDMKENPKCEIQKILK 209
QY 213 FLGVSCDKAQLAEALTEHCH-----QLVDQCCNAEALPVGRVGLMKDIFT 258
DB 210 FLEKDIPEILNKILYHSFSVMKGNPSANYTTMMKEEMDHVSFPMRKGISGDWKNQPT 269
QY 259 VSMNEKFDLVYKQKMGKCDLTF 280
DB 270 VAQYEKFEEDYVKRMEDSTLKF 291
RESULT 84
Q9Z1G0_RAT PRELIMINARY; PRT; 308 AA.
AC Q9Z1G0_RAT
DT 01-MAY-1999, integrated into UniProtKB/TrEMBL.
DT 01-MAY-1999, sequence version 1.
DT 07-FEB-2006, entry version 20.
DE Tyrosine-ester sulfotransferase (EC 2.8.2.9).
GN Name=Sult1d1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Sprague-Dawley; TISSUE=Kidney;

```
RA Herrmann A., Stoffel W.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; U32372; AAC99890.1; -; mRNA.
DR HSSP; P49888; 1H3.
DR Ensembl; ENSRNOG0000001960; Rattus norvegicus.
DR RGD; 620491; Sult1d1.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0017067; F:tyrosine-ester sulfotransferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
DR Transfaser.
SQ SEQUENCE 308 AA; 36341 MW; 68719C9223EE3EAE CRC64;

Query Match      26.9%; Score 412; DB 2; Length 308;
Best Local Similarity 32.3%; Pred. No. 6.1e-26;
Matches 93; Conservative 59; Mismatches 110; Indels 26; Gaps 5;

QY 15 PESKFEFHGRLPFCRGMKEIEANFVRPSDVWIVTPKSGTSLLOEVVYLVSGADP 74
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 8 FRRELVDVQGIPLFWSIAEQMSQVESFEARPDILISTYPKSGTTWISILDLYNNGDA 67
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 75 DEIGLMNIDQLVLVEYPOPL-----DIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIY 130
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 68 EKCKRDAIVRRVFFMELIIPGTINGVEMLDNMQSPRLVKTHLPVQLPSPSFWRNCKMIY 127
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 131 MARNEPKDLVSVYQFHRSLRTMSYGTQFCRCRFRMNDKLGYSWFEHVFQFWEHRMDSN 190
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 128 VARNAKDVAVSYFYHQAQKHPEPTWEEFLEKFMAGQVSPGPDYDHVKWGEKREYR 187
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 191 VLFLKYEDMHRDLVTMVQOLARFL-----GVSCDKAQLA-LTEHCHQL 233
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 198 ILICYEDMKEDPKCEIQKVLKLEKIDPEEVNVKILYHSSFSVMKANSANYTTMMKEE 247
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 234 VDQCNAALP-VGRGVGLWKDITVSNMKEFDLVYKQMGKCDLTP 280
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 248 MDQSVS----PPMRKGISGDMKNQFTVAQYKFEEDYVKMGESTLKF 291
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 85
Q9R2C2 MOUSE
ID Q9R2C2 MOUSE PRELIMINARY; PRT; 309 AA.
AC Q9R2C2;
DT 01-MAY-2000, integrated into UniProtKB/TrEMBL.
DT 01-MAY-2000, sequence version 1.
DT 07-FEB-2006, entry version 21.
DE Tyrosine-ester sulfotransferase (EC 2.8.2.9).
GN Name=Sult1d1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=BALB/c; TISSUE=Kidney;
RA Herrmann A., Stoffel W.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; U32371; AAC99889.1; -; mRNA.
DR HSSP; P49888; 1H3.
DR Ensembl; ENSMUSG00000029273; Mus musculus.
DR MGI; MGI:1926341; Sult1d1.
DR GO; GO:0004062; F:aryl sulfotransferase activity; IEA.
DR GO; GO:0000103; P:sulfate assimilation; IEA.
DR InterPro; IPR000863; Sulfotransferase.
```

```
DR Pfam; PF00685; Sulfotransferase; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
DR Transfaser.
SQ SEQUENCE 309 AA; 36713 MW; ESD5ELCD3C6D5117 CRC64;

Query Match      26.9%; Score 412; DB 2; Length 309;
Best Local Similarity 31.7%; Pred. No. 6.1e-26;
Matches 83; Conservative 63; Mismatches 98; Indels 18; Gaps 2;

QY 37 EIANFVRPSDVWIVTPKSGTSLLOEVVYLVSGADPDEIGLMNIDQLPVLVEYPOPL 96
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 30 QVESFEARPDILISTYPKSGTTWISILDLYNNGDAEKCKRDAIVKRVFPMELIIPGI 89
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 97 ----DIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPDLVSVYQFHRSLRTM 152
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 90 TNGVEMLNMPSPRIVKTHLPVQLPSPSFWKNDCKIIVARNAKDVVSYFYFYQMAKIH 149
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 153 SYRGTFQFCRCRFRMNDKLGYSWFEHVFQFWEHRMDSNVFLKYEDMHRDLVTMVQOLAR 212
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 150 PEPGTWEEFLEKFMAGQVSPGPDYDHVKWGEKREYRILYLFYEDMKENPKCEIQKILK 209
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 213 FLGVSCDKAQLA-LTEHCH-----QLVDQCNAALPVGGRGVGLWKDITP 258
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 210 FLEKIDPEILNKILYHSSFSVMKENSANYTTMMKEEMDSVSPFMRKGISGDMKNQFT 269
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 259 VSMNEKFDLVYKQMGKCDLTP 280
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 270 VAQYKFEEDYVKMGESTLKF 291
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 86
Q66KW4 XENLA
ID Q66KW4 XENLA PRELIMINARY; PRT; 305 AA.
AC Q66KW4;
DT 11-OCT-2004, integrated into UniProtKB/TrEMBL.
DT 11-OCT-2004, sequence version 1.
DT 07-FEB-2006, entry version 12.
DE MGC85375 protein.
GN Name=MGC8375;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaney S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
```

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Whole;
 RA Klein S., Gerhard D.S.;
 RA Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC -----
 DR EMBL; BC078536; AAH78536.1; -; mRNA.
 DR GO; GO:0008146; F:sulfotransferase activity; IEA.
 DR InterPro; IPR000863; Sulfotransferase.
 DR Pfam; PF00685; Sulfotransferase 1; 1.
 DR ProDom; PD001218; Sulfotransferase; 1.
 SQ SEQUENCE 305 AA; 35569 MW; PCS175F6S30564DF CRC64;
 Query Match 26.9%; Score 411; DB 2; Length 305;
 Best Local Similarity 33.7%; Pred. No. 7.3e-26;
 Matches 101; Conservative 48; Mismatches 123; Indels 28; Gaps 6;
 QY 1 MAESAEPTPTGFESEKVFEGHVRLP-PFCRGKMEIANPVPVDPVWVITYPKSGTS 59
 Db LAEEMENFQVTMGHTE-----GVLPQPTC-DEWDTIYFQAREDDILIATYPKSGTT 60
 QY 60 LLQEVVYLVSQAGDPEIGLNMIDQLPVLVYQPQ-----GLDIIKELTSPRLIKSHLPY 114
 Db WMQEIVDILQEGDVQKSWRAPCFIKVFIEMIPPKSMPSGLEGAKTWKSPRLIKTHLP 120
 QY 115 RFLPSDLHNGSKVIYMARNPDKLVVSYQFHRSLRTMSYRGTFQFCRRFNMNDKLGYS 174
 Db 121 NLLPFSFEKNAKVYVARNADQWVSYVYFHKMTFLDPCGTWDFSEFLSGDVPWGS 180
 QY 175 WFEHQVQFEHRMDSNVLFKYEDMHRDLVTVMEQLARPLGVSCDKAQLEALTEHCH--- 231
 Db 181 WFDHVLGVKAMDKHQILFIYEDMIEPFMRKVKMFLGDKLSDALENVKYHSSFQA 240
 QY 232 QLVDQCCNAEALP-----VGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTF 280
 Db 241 MKENPMTNNTVPSNIMDDTISPFMRKGLVGDWKTHFSVTQNFIDFKYKKQWEGSLNF 300
 RESULT 87
 ID Q5HZUO_XENTR PRELIMINARY; PRT; 297 AA.
 AC Q5HZUO;
 DT 15-FEB-2005, integrated into UniProtKB/TrEMBL.
 DT 15-FEB-2005, sequence version 1.
 DT 07-FEB-2006, entry version 10.
 DE Hypothetical LOC496998.
 GN Name=LOC496998;
 OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoides; Pipidae;
 OC Xenopodinae; Xenopus; Silurana.
 OX NCBI_TaxID=8364;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Whole body;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner B.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hong L.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshimiki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalios D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Whole body;
 RA Klein S., Gerhard D.S.;
 RA Submitted (JAN-2005) to the EMBL/GenBank/DBJ databases.
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC -----
 DR EMBL; BC088888; AAH88888.1; -; mRNA.
 DR GO; GO:0008146; F:sulfotransferase activity; IEA.
 DR InterPro; IPR000863; Sulfotransferase.
 DR Pfam; PF00685; Sulfotransferase 1; 1.
 DR ProDom; PD001218; Sulfotransferase; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 297 AA; 34863 MW; B8248162DC31D7D2 CRC64;
 Query Match 26.8%; Score 410.5; DB 2; Length 297;
 Best Local Similarity 32.4%; Pred. No. 7.7e-26;
 Matches 85; Conservative 61; Mismatches 99; Indels 17; Gaps 4;
 QY 34 KMEIANPVPVDPVWVITYPKSGTSLQEVVYLVSQAGDPEIGLNMIDQLPVLVY 92
 Db KIDSTQDFKVDVTVFLVYTKTQITQILSLIFNEGHRNGTEAIANV-PRVPIWY 88
 QY 93 QPGDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPDKLVVSYQFHRSLRTM 152
 Db 89 HSKVD-YDSRSPRLFSLSHLPYLPKDLRNKKKIIYVGRNPKDAVSYHYFNVIRL 147
 QY 153 SYRGTFQFCRRFNMNDKLGYSWFEHQVQFEHRMDSNVLFKYEDMHRDLVTVMEQLAR 212
 Db 148 KQVNDWESFLDYLTVGLGVSGWFDHVKGWYTHQEDFNILFVYEMKDLRSALVKICK 207
 QY 213 FLGVSCDKAQLEALTE-----HCHQLVDQCCNAEALFVGRGRVGLWKDIFT 258
 Db 208 FVEKELNEQVDITVERKATFKMKHDPPLANVTNSTDHLMKNGFTFLRRGTVDWKELMT 267
 QY 259 VSMNEKFDLVYKQMGKCDLTF 280
 Db VAQNEKFDKIYSEKMGVPINF 289
 RESULT 88
 ID STIE1_CAVPO STANDARD; PRT; 296 AA.
 AC P49887;
 DT 01-OCT-1996, integrated into UniProtKB/Swiss-Prot.
 DT 01-OCT-1996, sequence version 1.
 DT 07-FEB-2006, entry version 43.
 DE Estrogen sulfotransferase (EC 2.8.2.4) (Sulfotransferase, estrogen-
 preferring) (STIE3).
 DE Name=SULT1E1; Synonyms=STE;
 OS Cavia porcellus (Guinea pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 OC Hystricognathi; Caviidae; Cavia.
 OX NCBI_TaxID=10141;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [MRNA], AND PARTIAL PROTEIN SEQUENCE.
 RC STRAIN=NH 2; TISSUE=Adrenal cortex;
 RX MEDLINE=93024479; PubMed=1406700; DOI=10.1210/me.6.8.1216;
 RA Oeda T., Lee Y.C., Driscoll W.J., Chen H.-C., Strott C.A.;
 RT "Molecular cloning and expression of a full-length complementary DNA

```

RT encoding the guinea pig adrenocortical estrogen sulfotransferase.";
RL Mol. Endocrinol. 6:1216-1226(1992).
RN [2]
RP PAPS-BINDING SITE, AND MUTAGENESIS.
RA MEDLINE=95071451; PubMed=7980593;
RX Komatsu K., Driscoll W.J., Koh Y., Strott C.A.;
RT "A p-loop related motif (GxxGxxK) highly conserved in
RT sulfotransferases is required for binding the activated sulfate
RT donor.";
RL Biochem. Biophys. Res. Commun. 204:1178-1185(1994).
CC -!- FUNCTION: Sulfation of estradiol and estrone. May control the
CC level of the estrogen receptor by sulphylation free estradiol.
CC -!- CATALYTIC ACTIVITY: 3'-phosphoadenylyl sulfate + estrone =
CC adenosine 3',5'-biphosphate + estrone 3-sulfate.
CC -!- SUBUNIT: Homodimer.
CC -!- SUBCELLULAR LOCATION: Cytoplasm.
CC -!- TISSUE SPECIFICITY: Adrenal gland and much less in liver. Uterine
CC STE is detectable only during pregnancy.
CC -!- PTM: The N-terminus is blocked.
CC -!- SIMILARITY: Belongs to the sulfotransferase family.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; U09552; AAA18495.1; -; mRNA.
DR PIR; A44011; A44011.
DR SMR; P49888; IG3M.
DR HSP; P49887; 5-294.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase 1; 1.
DR ProDom; P0001218; Sulfotransferase; 1.
KW Direct protein sequencing; Lipid-binding; Steroid-binding;
KW Transferase.
FT CHAIN 1 296 Estrogen sulfotransferase.
FT /FTID=PRO_0000085152.
FT NP_BIND 49 54 PAPS (By similarity).
FT NP_BIND 131 139 PAPS (By similarity).
FT NP_BIND 194 230 PAPS (By similarity).
FT NP_BIND 238 260 PAPS (By similarity).
FT ACT_SITE 109 109 Proton acceptor (By similarity).
FT MUTAGEN 260 260 G->A: Loss of activity.
FT MUTAGEN 263 263 G->A: Loss of activity.
FT MUTAGEN 266 266 K->A: Loss of activity.
SQ SEQUENCE 296 AA; 35245 MW; 019D92895829901D CRC64;

Query Match 26.7%; Score 408.5; DB 1; Length 296;
Best Local Similarity 34.2%; Pred. No. 1.1e-25;
Matches 94; Conservative 50; Mismatches 112; Indels 19; Gaps 4;

QY 18 KYF-EFHGVLRLPPFCRGKMEETANFVRPSDVWIVTYPKSGTSLLEQVYLVYSGQADPDE 76
DB 11 EYFDFRGILLYKQIKYNDWNVEAFQAPDDDLVIAAPKSGTGTWSEVVCMIYASGDVKK 70

QY 77 IGLMNIDEOLPVLVEYPPQ-----GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMA 132
DB 71 CQDAIFENRVPLECENDKMWNGVQLEEMNSPRIIKTHLPRLLPASFWEKRCWKICIC 130

QY 133 RNPDLVSVYQFHRSLRTMSVYRGTFQEFRCFRFMDKLGYSQWGFHVFQEFWEHRMDSNVL 192
DB 131 RNAKDVAVSYYYFFLVMANHPDGSFPPEFVEKMGQVPGYSWYDHWKSVWMEKSTDPRIL 190

QY 193 FLKYEDMRDLVTMVEQLARFLGVSCDKAQLAEALTEHCH-----QLVDQCN 239
DB 191 FIFEDMKEDIRKVELKILHIFGRKPSBELVDKIHKTSFQEMKNPNSTNYTMLPEEIMN 250

QY 240 AEALP-VGRGRVGLWKDIFTVSMNEKFDLVYQKM 273
DB 251 QKVSPFMKGISGDMKQNHFTVALNESFDKHYQQM 285

RESULT 89
O70262_MOUSE PRELIMINARY; PRT; 304 AA.
ID O70262_MOUSE

```

```

AC O70262;
DT 01-AUG-1998, integrated into UniProtKB/TrEMBL.
DT 01-AUG-1998, sequence version 1.
DT 07-FEB-2006, entry version 26.
DE Phenol sulfotransferase.
GN Name=Sult1c1; Synonyms=Sult1a2;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ddv; TISSUE=Olfactory;
RX MEDLINE=96208491; PubMed=8641270;
RA Miyawaki A., Homma H., Tamura H., Matsui M., Mikoshiba K.;
RT "Zonal distribution of sulfotransferase for phenol in olfactory
RT sustentacular cells.";
RL EMBL J. 15:2050-2055(1996).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ddv; TISSUE=Olfactory;
RX MEDLINE=98228237; PubMed=9560327;
RA Tamura H.-O., Harada Y., Miyawaki A., Mikoshiba K., Matsui M.;
RT "Molecular cloning and expression of a cDNA encoding an olfactory-
RT specific mouse phenol sulphotransferase.";
RL Biochem. J. 331:953-958(1998).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ddv; TISSUE=Olfactory;
RA Tamura H.-O., Harada Y., Miyawaki A., Mikoshiba K., Matsui M.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF033653; AAC17740.1; -; mRNA.
DR HSP; P50224; 1CJM.
DR Ensembl; ENSMUSG0000003943; Mus musculus.
DR MGI; MGI:102928; Sult1c1.
DR GO; GO:0005737; Cytoplasm; IDA.
DR GO; GO:0004062; P-aryl sulfotransferase activity; IDA.
DR GO; GO:0006790; P-sulfur metabolism; IDA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase; 1.
DR ProDom; P0001218; Sulfotransferase; 1.
KW Transferase.
SQ SEQUENCE 304 AA; 35797 MW; 30857261A1462BE4 CRC64;

Query Match 26.7%; Score 408.5; DB 2; Length 304;
Best Local Similarity 31.5%; Pred. No. 1.2e-25;
Matches 88; Conservative 57; Mismatches 115; Indels 19; Gaps 3;

QY 21 EFHGVLRLPPFCRGKMEETANFVRPSDVWIVTYPKSGTSLLEQVYLVYSGQADPDEIGLM 80
DB 22 EVNGILMSKMNSWKNWQKLNWQAKPDDLLIATYAKAGITWTQEIIVDMIQNDGVDQKQRA 81

QY 81 NIDEOLPVLVEYPPQ-----GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNP 135
DB 82 NTDYRHPFIETWLPPLNSGLDLANKMPSRPTKTHLPVQMLPPSPFWKENSQIIYVARNA 141

QY 136 KDLVSVYQFHRSLRTMSVYRGTFQEFRCFRFMDKLGYSQWGFHVFQEFWEHRMDSNVLFK 195
DB 142 KDLVSVYQFHRSLRTMSVYRGTFQEFRCFRFMDKLGYSQWGFHVFQEFWEHRMDSNVLFK 201

QY 196 YEDMRDLVTMVEQLARFLGVSCDKAQLAEALTEHCHQLV---DQCCNAELP----- 244
DB 202 YEDMKEDPKREIKKIVKLEKDISSEVLNKKIHHSTFDMKQNPANVTTLTSSIMDSHI 261

QY 245 ---VGRGRVGLWKDIFTVSMNEKFDLVYQKMKGCDLTF 280
DB 262 SPFMKGMKPGDMKNYFTVAQSEDFEDYRKKMAGSTITF 300

```


Qy 245 VGRGRVGLWKDILFTVSMNEKEFDLVYKQKMGKCDLTF 280
Db 270 MRKGKGVDDWRNFTPEQEKQFEEDYKQKQKQVDPF 305

RESULT 91
Q9VHHO DROME PRELIMINARY; PRT; 316 AA.
ID Q9VHHO DROME PRELIMINARY; PRT; 316 AA.
AC Q9VHHO
DT 01-MAY-2000, integrated into UniProtKB/TrEMBL.
DT 01-MAY-2000, sequence version 1.
DT 07-FEB-2006, entry version 27.
DE CG16733-PA.
GN ORFNames=CG16733, Dmel CG16733;
OS *Drosophila melanogaster* (Fruit fly).
OC Eukaryota, Metazoa, Arthropoda, Hexapoda, Insecta, Pterygota;
OC Neoptera, Endopterygota, Diptera, Brachycera, Muscomorpha;
OC Ephydroidea, Drosophilidae, Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP NUCLOTIDE SEQUENCE.
RX MEDLINE=20196006; PubMed=107311132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Anantides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berland B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
RA Foster C., Gabriellat A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam K.A.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.C., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusekern D., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissensbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*."
RN Science 287:2185-2195(2000).
RN [2]
RP NUCLOTIDE SEQUENCE.
RX MEDLINE=2242665; PubMed=12533758;
RA Celniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Swirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstock R., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: release 3 of the *Drosophila*
RN melanogaster euchromatic genome sequence."
RN Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]

RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celisner S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
RT a genomics perspective";
RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Muncall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celisner S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yanada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RG Berkeley Drosophila Genome Project;
RA Celisner S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
RA Hoskins R., Stapleton M., Facleb J., Park S., Svirskas R., Smith E.,
RA Yu C., Rubin G.;
RT "Drosophila melanogaster release 4 sequence";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBSJ databases.
RN [6]
RP NUCLEOTIDE SEQUENCE.
RG Flybase;
RL Submitted (JAN-2006) to the EMBL/GenBank/DBSJ databases.
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AE003682; AAF54344.1; -; Genomic_DNA.
DR HSP; P49888; IHV3.
DR FlyBase; FBgn0037665; CG16733.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransfer 1; 1.
SQ SEQUENCE 316 AA; 37332 MW; EBCD9D561E4E1C77 CRC64;

Query Match 26.5%; Score 405; DB 2; Length 316;
Best Local Similarity 34.4%; Pred. No. 2.4e-25;
Matches 94; Conservative 52; Mismatches 99; Indels 28; Gaps 8;

QY 36 EBIANFPVPSVWIVTYPKSGTSLQEVVYLSQ-----GADPD-----EI-GLMN 81
DB 45 ESIRSLPVYQDDVMWVSYPRTGQWAEVWVLGHQLDYVAAEQDLRLRLSPILSALFS 104

QY 82 IDEQLPVLEYPOGLDIIKELTSRLIKSHLPVRLPSDLHNGDSKVIMARNPKDLVVS 141
DB 105 IDHETVTAQFGNTVDLVNLRPRFARSHLPWLLPEQFETVKPRIVTARNPKDLVVS 164

QY 142 YYQFHSRLTMSYRGTFQFCRRFMNDKLGYSWFHVEQFWEHRMDSNVLFKYEDMHR 201
DB 165 VYHYFKLLHGMN--GDPEQFVDFLEIGHTPMGSYWRHVLPPFKRSQDDNVLFKYEDVVK 222

QY 202 DLVTVWEQLARFLGVS--CDKAQLALTEHC---HQLVDCCNAE-ALP-----VGRGR 249
DB 223 DLPSVVRRCARFLGVSQSLDVSFTLQKLDHLLTFDKMRANKANLEKLLPSESSKFIIRNGK 282

QY 250 VGLWKDIFTVSMNEKFDLVYKMGKCDLTFDF 282
DB 283 IGDRNHNHGMNSERFDEWTERHMRGSLNFDY 315

RESULT 92
Q6DJL6_XENLA PRELIMINARY; PRT; 287 AA.
ID Q6DJL6_XENLA

AC O6DJL6;
DT 16-AUG-2004, integrated into UniProtKB/TrEMBL.
DT 16-AUG-2004, sequence version 1.
DT 07-FEB-2006, entry version 14.
DE MGC82053 protein.
GN Name=MGC82053;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=83355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Boak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,
FAhey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanches A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RA Klein S., Gerhard D.S.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBSJ databases.
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BC075160; AAH75160.1; -; mRNA.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR InterPro; IPR002202; HMG-CoA_red.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransfer 1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
DR PROSITE; PS01192; HMG COA REDUCTASE 3; UNKNOWN 1.
SQ SEQUENCE 287 AA; 33812 MW; EB655BD8D29B0AB7 CRC64;

Query Match 26.4%; Score 404.5; DB 2; Length 287;
Best Local Similarity 30.9%; Pred. No. 2.4e-25;
Matches 85; Conservative 57; Mismatches 116; Indels 17; Gaps 3;

QY 23 HGVLLPFCRCGMIEIANFPVPSVWIVTYPKSGTSLQEVVYLSQADPDEIGLMI 82
DB 9 NGVPLSEAIASNDWEIWFQAKPGDILINTPKSGTTWQELIDIMNDVDEKCRAPI 68

QY 83 DEQLP---VLEYPOGLDIIKELTSRLIKSHLPVRLPSDLHNGDSKVIMARNPKDLV 139
DB 69 YERIPFVDILHLMKPGVEEVNAMPSPRLKSHLAFQLVPPSPFWKCKVYVARNAKDTA 128

```

Qy      140  VSYQYHRSRLRTWSVRGTFQECRFRPMNDKLGYSWFEEHVOEFWEHHRMDSNVFLFKYEDM 199
Db      129  TSFFYPDHAHLHPTPEGWDDYLERFMKGDYGVGSWTDYKVGFEQKQDNILYLFYEDL 188
Qy      200  HRDLVTMVQLARFLGVSCDKAQLEA---LTEHCHQLVDQCCNAALP-----V 245
Db      189  QKNPLVEIRKVMKFLDKOLPEMLRKIVHLTSFKKMKENPMANYSTFSPDMLKQTDHKPM 248
Qy      246  GRGRVGLWKDIFTVSNEKFDLVYKQKMGKCDLTF 280
Db      249  RKGKVDGNGNFTVYONEVFADYOROMSGSTLKF 283

```

RESULT 93	
Q5M7C2_XENLA	PRELIMINARY; PRT; 287 AA.
ID Q5M7C2	
AC Q5M7C2	
DT 01-FEB-2005	integrated into UniProtKB/TrEMBL.
DT 01-FEB-2005	sequence version 1.
DT 07-FEB-2006	entry version 10.
DE LOC496246	protein.
DE Name=LOC496246;	
GN Xenopus laevis (African clawed frog).	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;	
OC Xenopodinae; Xenopus; Xenopus.	
OX NCBI_TaxID=8355;	
RN [1]	
RP NUCLEOTIDE SEQUENCE.	
RC TISSUE=Testis;	
RX MEDLINE=223411132; PubMed=124549417; DOI=10.1002/dvdy.10174;	
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,	
RA Richardson P.,	
RA "genetic and genomic tools for Xenopus research: The NIH Xenopus	
RT initiative";	
RT Dev. Dyn. 225:384-391(2002).	
RN [2]	
RP NUCLEOTIDE SEQUENCE.	
RC TISSUE=Testis;	
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;	
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,	
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,	
RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,	
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,	
RA Diatchenko L., Marusina K., Farmer A.A., Rubin A.A., Hong L.,	
RA Sapletton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,	
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,	
RA Raha S.S., Loquellano N.A., Peters G.J., Malek J.A., Gunaratne P.H.,	
RA Bosak S.A., McSwain P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,	
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,	
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,	
RA Fahey J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A.,	
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,	
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,	
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,	
RA Butterfield F.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,	
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;	
RT "Generation and initial analysis of more than 15,000 full-length human	
RT and mouse cDNA sequences";	
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).	
RN [3]	
RP NUCLEOTIDE SEQUENCE.	
RC TISSUE=Testis;	
RA Klein S., Gerhard D.S.;	
RA Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.	
CC	
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms	
CC Distributed under the Creative Commons Attribution-NoDerivs License	
CC	
CC EMBL; BC088717; AAH88717.1; -; mRNA.	
DR GO; GO:0008146; F:sulfotransferase activity; IEA.	
DR InterPro; IPR000863; Sulfotransferase.	

[illegible]

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 17, 2006, 11:17:09 ; Search time 10 Seconds
(without alignments)
60.666 Million cell updates/sec

Title: US-10-768-158-2
Perfect score: 1530
Sequence: 1 MAESEATPTGFESEKYF.....FDLVYQKMGKDLTFDFYL 284

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 21570 seqs, 2136119 residues
Total number of hits satisfying chosen parameters: 21570

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : Published Applications AA.New:*
1: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
2: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
3: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
4: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
5: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
6: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
7: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
8: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	77.5	5.1	411	6	US-10-505-928-445
2	75	4.9	337	6	US-10-196-749-268
3	75	4.9	337	7	US-11-101-316-74
4	73.5	4.8	273	6	US-10-511-814-1
5	72.5	4.7	365	7	US-11-251-465-25
6	72.5	4.7	367	7	US-11-312-958-14
7	72	4.7	813	6	US-10-196-749-466
8	71	4.6	259	7	US-11-249-111-118
9	69	4.5	1534	6	US-10-505-928-40
10	68.5	4.5	634	7	US-11-311-778-18
11	68.5	4.5	820	6	US-10-982-908-26
12	67.5	4.4	424	7	US-11-242-505A-33
13	67.5	4.4	1056	6	US-10-505-928-225
14	65	4.2	587	6	US-10-511-937-2407
15	64.5	4.2	2202	6	US-10-488-015-12
16	64	4.2	226	7	US-11-183-218-42
17	63.5	4.2	209	6	US-10-196-749-270
18	63.5	4.2	226	6	US-10-984-694-3
19	62.5	4.1	226	6	US-10-984-694-1
20	62.5	4.1	933	6	US-10-511-455-50
21	62.5	4.1	956	6	US-10-511-455-41
22	62	4.1	309	7	US-11-283-522-34
23	62	4.1	492	6	US-10-505-928-642
24	62	4.1	1085	6	US-10-505-928-343
25	62	4.1	1722	6	US-10-505-928-780
26	61.5	4.0	226	6	US-10-984-694-5
27	61.5	4.0	512	6	US-10-505-928-233
28	61.5	4.0	789	6	US-10-511-937-2512
29	61	4.0	239	7	US-11-211-917-40
30	60.5	4.0	811	6	US-10-511-937-2584
31	60.5	4.0	3020	7	US-11-140-487A-772
32	60	3.9	1457	7	US-11-280-757-37
33	60	3.9	2351	7	US-11-183-218-30
34	60	3.9	2351	7	US-11-280-757-35
35	59.5	3.9	310	7	US-11-101-316-140
36	59.5	3.9	347	6	US-10-196-749-404
37	59.5	3.9	513	7	US-11-101-316-124
38	59.5	3.9	694	7	US-11-252-080-2
39	59.5	3.9	1807	6	US-10-504-120-22
40	59	3.9	734	7	US-11-238-282-19
41	59	3.9	1085	6	US-10-505-928-175
42	59	3.9	578	6	US-10-505-928-401
43	58.5	3.8	630	7	US-11-146-700-3
44	58.5	3.8	727	6	US-10-505-928-644
45	58.5	3.8	727	6	US-10-511-937-2606
46	58.5	3.8	763	6	US-10-505-928-304
47	58.5	3.8	3396	6	US-10-505-928-449
48	58.5	3.8	458	6	US-10-196-749-496
49	58	3.8	1005	6	US-10-473-691B-2
50	58	3.8	1504	6	US-10-505-928-662
51	57.5	3.8	308	7	US-11-251-466-2
52	57.5	3.8	609	6	US-10-505-928-646
53	57.5	3.8	708	6	US-10-505-928-416
54	57.5	3.8	1763	6	US-10-504-120-21
55	57	3.7	720	6	US-10-505-928-288
56	57	3.7	832	6	US-10-505-928-491
57	57	3.7	926	6	US-10-505-928-35
58	57	3.7	2871	6	US-10-505-928-100
59	57	3.7	493	6	US-10-505-928-167
60	56.5	3.7	505	6	US-10-511-937-2991
61	56.5	3.7	642	7	US-11-301-924-4
62	56.5	3.7	147	6	US-10-468-193-7
63	56	3.7	275	7	US-11-264-784-54
64	56	3.7	450	6	US-10-489-730-4
65	56	3.7	459	6	US-10-528-032-17
66	56	3.7	480	7	US-11-301-554-336
67	56	3.7	578	1	US-09-949-925-82
68	56	3.7	587	6	US-10-489-730-2
69	56	3.7	661	6	US-10-489-730-11
70	56	3.7	736	6	US-10-511-455-56
71	56	3.7	736	7	US-11-219-635-3
72	56	3.7	736	7	US-11-254-185-8
73	55.5	3.6	230	7	US-11-254-185-40
74	55.5	3.6	230	7	US-10-370-959-145
75	55.5	3.6	448	6	US-11-249-111-84
76	55.5	3.6	448	7	US-10-505-928-803
77	55.5	3.6	535	6	US-10-505-928-393
78	55.5	3.6	667	6	US-10-534-462-1
79	55.5	3.6	847	6	US-10-497-088-21
80	55.5	3.6	1052	6	US-10-505-928-729
81	55.5	3.6	1060	6	US-10-497-088-14
82	55.5	3.6	1342	7	US-11-301-554-1903
83	55.5	3.6	2209	7	US-10-511-937-2546
84	55	3.6	274	6	US-11-283-522-35
85	55	3.6	300	7	US-10-511-937-2511
86	55	3.6	343	6	US-11-305-413-13
87	55	3.6	378	7	US-10-196-749-152
88	55	3.6	541	6	US-10-511-937-2413
89	55	3.6	750	6	US-11-266-747-1390
90	54.5	3.6	254	7	US-10-538-066-366
91	54.5	3.6	314	6	US-10-505-928-741
92	54.5	3.6	374	6	US-10-510-325-12
93	54.5	3.6	511	6	US-11-301-554-1819
94	54.5	3.6	831	6	US-10-505-928-359
95	54.5	3.6	1015	6	US-11-311-778-24
96	54.5	3.6	1177	7	US-11-311-778-10
97	54.5	3.6	1178	7	US-11-311-778-12
98	54.5	3.6	1178	7	US-11-311-778-12

99 54.5 3.6 1178 7 US-11-311-778-14 Sequence 14, Appl
100 54.5 3.6 1575 6 US-10-505-928-257 Sequence 257, App

ALIGNMENTS

RESULT 1

US-10-505-928-445
; Sequence 445, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 445
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-445

Query Match 5.1%; Score 77.5; DB 6; Length 411;
Best Local Similarity 17.3%; Pred. No. 0.5;
Matches 51; Conservative 56; Mismatches 84; Indels 103; Gaps 13;
QY 19 YPEFHGVRLLP-----FCGKMEETANFPV-----RPSDVWIVTPKSGTSLLEQVYV 66
Db 133 YPLENIKPPVNNHTTDRIFRERGASRVLCSPVCPDPPGADL----- 174
QY 67 LVSQADPDEIGLMNID-----EQLPVLEYQPQGLDIKELT-SPRLIKSHLPYRF 116
Db 175 VLEEGDCVRKCGLLNLTVAAEACRERSHVAIKTVRVPENDURALVEDPRL----- 225
QY 117 LPSDLHGDSKVIYMARNPDKDLVSVYQFHRSLRTMSVGTGTFQCFRRFMDKLGYG--- 173
Db 226 -----NLKVIQLVRDPGILAS-----RSETFRDTYRLM-----RLWYGTGR 262
QY 174 -----SWFEHVQFWEHRMDSNV-----LFLKYEDMHRDLVTMVEQLARFLGV 216
Db 263 KPNLDVTQLTTVCDFNSVSTGLMRPWLKGYMLVRYEDLARNPMKKTBEIYGFLGI 322
QY 217 SCDKAQLREALTEHQVLVDQCCNAEALP-VGRGRVGLWKDIFTVSMNEKFDLVY 269
Db 323 PLD-----SHVARWION--NTRGDDTLGKHKGTVRNSAATAEKWRFLSY 366

RESULT 2

US-10-196-749-268
; Sequence 268, Application US/10196749
; Publication No. US20060094864A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C340
; CURRENT APPLICATION NUMBER: US/10/196,749
; CURRENT FILING DATE: 2002-07-16

; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 268
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-196-749-268

Query Match 4.9%; Score 75; DB 6; Length 337;
Best Local Similarity 18.8%; Pred. No. 0.7;
Matches 33; Conservative 16; Mismatches 41; Indels 86; Gaps 7;
QY 47 DWIVTTPKSGTSLLEQV-VVLVSQGADPPDEIGLMNIDTEQLPVLE---YPOPGLDIIKEL 102
Db 31 DHPFASYPECGNNAQSPIDIQTDSVTFDPD-----LPALQPHGYDQPGTE----- 75
QY 103 TSPRLIKSHLPYRFLPDLHGDSKV-----TYMARNPDKDLVSVYQFHRSLRTMSYR 155
Db 76 -----PLDHNHGHVTVQLSLPSTLYLGLPKRYVAAQLHLH----- 111
QY 156 GTFQBFCCRFRMNDKLGYSWFEHVQEFW-----EHRMDSNVFLFKYEDMHRD 202
Db 112 -----WGKGSFGSGSEHQINSEATFAELHIVHYD 140

RESULT 3

US-11-101-316-74
; Sequence 74, Application US/1101316
; Publication No. US20060099657A1
; GENERAL INFORMATION:
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: ANTIBODIES TO A POLYPEPTIDE ENCODED BY A NUCLEIC ACID
; FILE REFERENCE: P3230R1C17C1
; CURRENT APPLICATION NUMBER: US/11/101,316
; CURRENT FILING DATE: 2005-04-06
; PRIOR APPLICATION NUMBER: 10/063526
; PRIOR FILING DATE: 2002-05-03
; PRIOR APPLICATION NUMBER: 10/006867
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 09/380137
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: PCT/US99/12252
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; NUMBER OF SEQ ID NOS: 170


```
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 60/373,010
; PRIOR FILING DATE: 2002-04-16
; PRIOR APPLICATION NUMBER: US 60/373,908
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/377,717
; PRIOR FILING DATE: 2002-05-03
; PRIOR APPLICATION NUMBER: US 60/379,949
; PRIOR FILING DATE: 2002-05-13
; PRIOR APPLICATION NUMBER: US 60/382,409
; PRIOR FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: US 60/385,280
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 60/386,879
; PRIOR FILING DATE: 2002-06-06
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-196-749-466
```

```
Query Match      4.7%; Score 72.5; DB 7; Length 367;
Best Local Similarity 21.8%; Pred. No. 1.5; Mismatches 35; Indels 43; Gaps 9;
Matches 43; Conservative 35; Mismatches 35; Indels 43; Gaps 9;

Qy 54 PKSGTSL- QEVVYVLSQADPDEIGLMNIDQLPVLEYPQP-----GLDIIKELT 103
Db 107 PKLGTSLPQALIVGVKGGTRAVLEFIRVHPDVRAIG-TEPHFPDRNYGRGLDWYRSL- 164

Qy 104 SPRLIKSHLPYRFLPSDL-----HNGDSKVIYMARPKDLVVSYYQFHRSLRTM 152
Db 165 MERTLESQITTEKTSYFVTOEAPRIFNMRDTRKLIIVVRNPVTRAIISDY-----TOTL 219

Qy 153 SYR---GTFQSFRRFMNDKLG-----YGSWFHVOEFWEHRMDSNVFLFKYEDM 199
Db 220 SKKPDIPFEGI--SFRNRTGLVDVSNVNAIRIGMYVHLESWLQYFFLAQIHVFSGERL 277

Qy 200 HRDLVTMVEQLARFLGV 216
Db 278 ITDPAGEMGRVQDFLGI 294
```

```
RESULT 7
US-10-196-749-466
; Sequence 466, Application US/10196749
; Publication No. US20060094864A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430RIC340
; CURRENT APPLICATION NUMBER: US/10/196,749
; CURRENT FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
```

```
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 466
; LENGTH: 813
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-196-749-466
```

```
Query Match      4.7%; Score 72; DB 6; Length 813;
Best Local Similarity 21.3%; Pred. No. 5.8;
Matches 60; Conservative 37; Mismatches 85; Indels 100; Gaps 15;

Qy 42 PVRPDSVWITYPKSGTSLQEVVYVLSQADPDEIGLMNIDQLPVLEYPQPGLDIIKE 101
Db 39 PVTPH--WVLDGGQWRTVSLEEPV-----SKPD-MGLVALEAE-----QCELULE 80

Qy 102 -----LTSPLRIKSHL-----PYRFLPSDLH-----NGDSKVI 129
Db 81 LEKNHRLLAGYIETHYGPDPQPVVLPAPNHTDCHYQGRVGRPDSNVVLTCSGMSGLI 140

Qy 130 YMARNPDKLVVSY-----OFHRSLSRTMSYRGTFQEPFC-RRFNMNDKLGYG 173
Db 141 TLSRN-----ASYVLRPWPPRGSKDFSTHEIFRMEQLLTWKGT-----CGHRDPGNKAGMT 191

Qy 174 SW-----PEHVQEFWE-HRMSNVFLFKYEDMHRDL-----VTMVQOLA 211
Db 192 SLPGGSGRGRREARRTKYKLYELIVADHTLFL---TTRHNLNHTKORLLELVANYVDQLL 249

Qy 212 RFLGVSCDKAQLAETECHQLVDQCCNAEALPVGRGRVGLW 253
Db 249 RTLDIQVALTGLEVWTERDSRSRVTDANATILWAPLOWRGLW 290
```

```
RESULT 8
US-11-249-111-118
; Sequence 118, Application US/11249111
; Publication No. US20060099623A1
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Lubbers, Mark W
; APPLICANT: Dekker, James
; TITLE OF INVENTION: Polynucleotides and polypeptides isolated from Lactobacillus
; FILE REFERENCE: 13353.1048uic2
; CURRENT APPLICATION NUMBER: US/11/249,111
; CURRENT FILING DATE: 2005-10-11
; PRIOR APPLICATION NUMBER: 10/288,930
; PRIOR FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: 09/724,623
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 60/148,801
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 124
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 118
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Lactobacillus rhamnosus
US-11-249-111-118
```

Query Match 4.6%; Score 71; DB 7; Length 259;
Best Local Similarity 24.6%; Pred. No. 1.3;
Matches 32; Conservative 18; Mismatches 44; Indels 36; Gaps 7;

QY 54 PKSGTSLQEVVYLSQGADPD-----EIGLMNIDQLPVLVYPPQGLDII 99
DB 66 PEEATELLWEAIFKTRIDQDPPEAAWKAHQKLEKAAWLN-NEQFDOLHYMAFGTDLV 124
QY 100 KELTSPRLIKSHL-----PYRFLPSDLHNGDSKVIYMARPNPKOLVVSYYOFHR-----SLR 150
DB 125 VGLP-----KNHWEAGAFNP-----RGE-----FMANMPTBEVFTAPDFRRIDGTGVAATK 172
QY 151 TMSYRGTFQE 160
DB 173 PLSYGGNILE 182

RESULT 9

US-10-505-928-40
; Sequence 40, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR FILING DATE: 2004-03-07
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: Patent in 3.2
; SEQ ID NO 40
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-40

Query Match 4.5%; Score 69; DB 6; Length 1534;
Best Local Similarity 22.1%; Pred. No. 33;
Matches 55; Conservative 39; Mismatches 73; Indels 82; Gaps 14;

QY 91 YPQGLDIIKELTSRLIKSHLPYRFLPSDLHNGDS-KVIYMAR--NPKDL----- 138
DB 1282 YKEPKLTPLSEI-SORLLK-----LYSDKPGSENKVMQDSKGNKPNKDLDSKVAYIQV 1333
QY 139 --VVSYY-----QFHR-----LRTMSYRGTFQFCRRFMNDKLGYS 174
DB 1334 THVIFPFDEKELQERKTEFERSHNIRRFMPFPTQTKRQGVVEQCKR--RTILTAIH 1391
QY 175 WFEHQE-----FWEHRMDSNVLFLKYEDMHRDLVTMVEQLARFLGVSCDKAQLAETHC 230
DB 1392 CFPYVKRIPVMYQHHTDLNPIEVAIDEMSKK-VAE LRQLC SSAEVDIMKQLKLQGSVS 1450
QY 231 HOLVQCCNAEALPVR-----GRVGLWKDIF-----TVSMNEKF--- 265
DB 1451 VQV-----NAGFLAYARAFLODNTNRYPDNKKVLLKEVFQVBEACQALAVNERLIKE 1505
QY 266 -DLVYKQKM 273
DB 1506 DQLEYQEM 1514

RESULT 10

US-11-311-778-18
; Sequence 18, Application US/11311778
; Publication No. US20060095988A1
; GENERAL INFORMATION:
; APPLICANT: Fischhoff, David A.
; APPLICANT: Perlak, Frederick J.
; TITLE OF INVENTION: SYNTHETIC PLANT GENES AND METHOD FOR PREPARATION
; FILE REFERENCE: 38-21 (13553) C
; CURRENT APPLICATION NUMBER: US/11/311,778
; CURRENT FILING DATE: 2005-12-19

; PRIOR APPLICATION NUMBER: 08/434,105
; PRIOR FILING DATE: 1995-05-03
; PRIOR APPLICATION NUMBER: 07/959,506
; PRIOR FILING DATE: 1992-10-09
; PRIOR APPLICATION NUMBER: 07/476,661
; PRIOR FILING DATE: 1990-02-12
; PRIOR APPLICATION NUMBER: 07/315,355
; PRIOR FILING DATE: 1989-02-24
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 18
; LENGTH: 634
; TYPE: PRT
; ORGANISM: artificial
; FEATURE:
; OTHER INFORMATION: synthetic structural gene encoding P2 insecticidal protein
US-11-311-778-18

Query Match 4.5%; Score 68.5; DB 7; Length 634;
Best Local Similarity 19.4%; Pred. No. 9.6;
Matches 45; Conservative 42; Mismatches 78; Indels 67; Gaps 13;

QY 46 SDWIVTVPKSGTSLQEVV-----YLVSQGADPDPEIGLMN-----IDEQ 85
DB 76 SELWGIIFPSGSTNLMDILRETEQFLNQLRLTDTLARVNAELIGLQANIREFNQVDNF 135
QY 86 LPVLEYQPGLDIIKEL-TSPRLIKSHLP-----YRFL-----PSDLH----- 122
DB 136 LNPTQNPVP-LSITSSVNTMQQLFLNRLPQFQIQYQLLLPLFAQAANMHLSFIRDVIL 194
QY 123 NGDSKVIYMA--RNPKDLVVSYYQFHRSLRTMSYRGTFQECRBFMDKLGYSW----- 175
DB 195 NADEWGISAATLRTYRDYLRNYTRDYSNYCINTYTQAFRGLNTR-LHDMLEFRYMFNLV 253
QY 176 FEHQVEFWEHRMDSNVLFLKYEDMHRDLVTMVEQLARFLGVSCDKAQLAALT 227
DB 254 FEYV-SIWS-----LFKYQSL-----MVSSGANLYASGSGPQQTQST 290

RESULT 11

US-10-982-908-26
; Sequence 26, Application US/10982908
; Publication No. US2006009603A1
; GENERAL INFORMATION:
; APPLICANT: DAVI, FREDERIC
; APPLICANT: DIGHERIO, GUILLAUME
; APPLICANT: DUMAS, GERARD
; APPLICANT: OPPEZO, PABLO
; APPLICANT: SETTEGRANA, CATHERINE
; APPLICANT: VASCONCELOS PINHEIRO, YURI
; TITLE OF INVENTION: METHOD OF DIAGNOSIS/PROGNOSIS OF HUMAN CHRONIC LYMPHOCYTIC
; FILE REFERENCE: 259358USO
; CURRENT APPLICATION NUMBER: US/10/982,908
; CURRENT FILING DATE: 2004-11-08
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 26
; LENGTH: 820
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-982-908-26

Query Match 4.5%; Score 68.5; DB 6; Length 820;
Best Local Similarity 19.2%; Pred. No. 14;
Matches 71; Conservative 42; Mismatches 81; Indels 175; Gaps 21;

QY 31 CRGKMEETANFPVRPSDV---WI-----VTYPKSGTSLQEVVYLSQG 71
DB 173 CRMEFEEDNSTQKQSSVYGVGWHIFRIVEIVVVDNVLVIRYERNDKSLLED-LYIVNI 231
QY 72 ADP--DEIGLMNIDEQLPVLEYQPGLDIIKELTSPLRI-----KS-HLPYRFLPSDLH 122

```
Db 232 VDSILDVIG-----VKVLLF---GLEI---WTNKNLIVDDVRKSVHLYCKWKSENI- 277
Qy 123 NGDSKVIYARNPKDLVVSYVQFHRSLRTMSYRGTFQBFQR-----RFMNDKLG--- 171
Db 278 -----TPRMQHD--TSHLFTTLGLRGLSGIGAFRGKCTPHRSCAIVTFMNTKLTGTF 327
Qy 172 -----VGSFHFVQBEWEHRMD 188
Db 328 IAVAHHLGHNIGMNHDEDTCSQPCIMHGNPPITKPSNCSYG-----DFWEYTYE 380
Qy 189 SNVLFLKYEDMHRDLVTMVEQLAREL---GVSCDKAQLEALTEHCHQLVDOCC----- 238
Db 381 RTKCLL--ETVHTKDFVVKRCGVGVBEBCDGPL-----KHCAK--DFCCLSNCTLT 432
Qy 239 -----NARALPVGRRVGLWKDIFTVSMNEKFDLVYKQKMGKCDL----- 278
Db 433 DOSTCAFLGCKCKDFLPSGK-----VCRKEVNECDLPENCNGTS 472
Qy 279 ---TFDFYL 284
Db 473 HKCPDFFVY 481

RESULT 12
US-11-242-505A-33
; Sequence 33, Application US/11242505A
; Publication No. US20060099656A1
; GENERAL INFORMATION:
; APPLICANT: Carroll, Joseph M.
; TITLE OF INVENTION: Methods and Compositions for Treating
; TITLE OF INVENTION: Hematological Disorders Using 232, 2059, 10630, 12848, 13875,
; FILE OF INVENTION: 14395, 14618, 17692, 58874, 252, 304, 1980, 14717, 9941, 19310,
; FILE REFERENCE: MPI2001-288PRCPIONIM
; CURRENT APPLICATION NUMBER: US/11/242.505A
; CURRENT FILING DATE: 2005-10-03
; PRIOR APPLICATION NUMBER: US 10/290,078
; PRIOR FILING DATE: 2002-11-07
; PRIOR APPLICATION NUMBER: US 60/347,949
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 10/320,351
; PRIOR FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: 60/341,606
; PRIOR FILING DATE: 2001-12-17
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-242-505A-33

Query Match 4.4%; Score 67.5; DB 7; Length 424;
Best Local Similarity 30.3%; Pred. No. 6.7;
Matches 27; Conservative 13; Mismatches 40; Indels 9; Gaps 4;

Qy 38 INFVPRSDWIVTPKSGTSL--QEVVYL--VSGADPDDEIGLMNIDEQL-----PVLE 90
Db 83 VALFQVLPOLLIDITYRFQGPDLRCRAVKYQLQVLSNFASTYMLLANTLDRVLAVCHPLRS 142
Qy 91 YPQPGLDIIEKETSRLIKS--HLPVREL 117
Db 143 LQPGQSTVLLIAPWLLAAIFSLQVFI 171

RESULT 13
US-10-505-928-225
; Sequence 225, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178

; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 225
; LENGTH: 1056
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-225

Query Match 4.4%; Score 67.5; DB 6; Length 1056;
Best Local Similarity 18.2%; Pred. No. 27;
Matches 52; Conservative 43; Mismatches 92; Indels 99; Gaps 11;

Qy 35 MEEIANFPVRPSDWMIVTPKSGTSL--QEVVYLVSQADPDDEIGLMNIDEQLPVLEYQPQ 94
Db 303 VERTPHVPYRESKLRILQDSLGRTRTSIATISPAS-----LNLEETLSTLEYAHR 355
Qy 95 GLDII-----KELTSPRLIKSHL--PYRFLPSDLHNGDSKVIYARNPKDLVVSYVQFHR 147
Db 356 AKNILNKPENQKLTKKALIKETTERLKRDL-----AAAREKNGVYIISEENF-- 405
Qy 148 SLRTMSYRGTFQE-----FCRRFMNDKLGYSW 175
Db 406 --RVMSGKLTVOBEQIVELTEKICAVEEELNRVTELFMDNKNELDOCKSDLNKNT----- 458
Qy 176 FEHVQEFWEHRMDSNVLFKYE-----DMHRDLVTMVEQLARFL--GVSCDK 220
Db 459 -QLETTQCHLQETKQLVKEEYITSALSTEELKHLDAASKLLNTVETTTKDVSGLHSL 517
Qy 221 AQLBALTEHCHQLVDQCNARALPVGRRVGLWKDIFTVSMNEKFD 266
Db 518 DRKKAVDQH-----NABA-----QDIFGKNLSLNF 543

RESULT 14
US-10-511-937-2407
; Sequence 2407, Application US/10511937
; Publication No. US20060088836A1
; GENERAL INFORMATION:
; APPLICANT: EXPRESSION DIAGNOSTICS, INC.
; APPLICANT: Wohlgenuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; APPLICANT: Prentice, James
; APPLICANT: Morris, MacDonald
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND MONITORING TRANSPLANT REJECTION
; FILE REFERENCE: 506612000104
; CURRENT APPLICATION NUMBER: US/10/511,937
; CURRENT FILING DATE: 2004-10-19
; PRIOR APPLICATION NUMBER: PCT/US2003/012946
; PRIOR FILING DATE: 2003-04-24
; PRIOR APPLICATION NUMBER: US 10/131,831
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US 10/325,899
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 3117
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2407
; LENGTH: 587
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-511-937-2407

Query Match 4.2%; Score 65; DB 6; Length 587;
Best Local Similarity 24.7%; Pred. No. 21;
Matches 42; Conservative 24; Mismatches 56; Indels 48; Gaps 9;
```



```
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: GB0207904.4
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: GB0207900.2
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: GB0207901.0
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: GB0227734.1
; PRIOR FILING DATE: 2002-11-28
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 50
; LENGTH: 933
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-511-455-50

Query Match
Best Local Similarity 4.1%; Score 62.5; DB 6; Length 933;
Matches 47; Conservative 28; Mismatches 62; Indels 107; Gaps 10;

QY 26 RLPPFCRGMBEIA-----NFPVRPSDVWVITPKSGTSLLOEVVYLVVSQGADPDEIGLM 80
Db 420 RYEGFCVDMKELAEILRFNYKIRLVGDGVGYVPEANGTWTGNGELIARKADLAVAGLT 479
QY 81 NIDEQLPVLEYPOP-----GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYARNPK 136
Db 480 ITAREKVIDFSKPFMTLGISIL-----YR-----IHMGRKP- 511
QY 137 DLVSVYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVFQEFWEHRMDSNVLFKY 196
Db 512 ----GYFSF-----LDPFSPGW-----LFM-- 528
QY 197 EDMHRDLVTMVEQLARFLGVSCDKAQLAEALTEH-----CHQLVDQCCNAEAL- 243
Db 529 -----LLAYLAVSCVLFVARLTPYEWYSPHPCAQGRCNLLVNGYSLGSLW 575
QY 244 -PVG 246
Db 576 FPVG 579

RESULT 21
US-10-511-455-41
; Sequence 41, Application US/10511455
; Publication No. US2006008835A1
; GENERAL INFORMATION:
; APPLICANT: Pickard, Benjamin Simon
; APPLICANT: Blackwood, Douglas
; APPLICANT: Porteous, David
; APPLICANT: Muir, Walter John
; APPLICANT: Mors, Ole
; APPLICANT: Ewald, Henrik Lykke
; TITLE OF INVENTION: SCHIZOPHRENIA ASSOCIATED GENES
; FILE REFERENCE: 9013.63
; CURRENT APPLICATION NUMBER: US/10/511,455
; PRIOR FILING DATE: 2004-10-05
; PRIOR APPLICATION NUMBER: PCT/GB03/001543
; PRIOR FILING DATE: 2003-04-07
; PRIOR APPLICATION NUMBER: GB0207902.8
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: GB0207904.4
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: GB0207900.2
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: GB0207901.0
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: GB0227734.1
; PRIOR FILING DATE: 2002-11-28
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 41
; LENGTH: 956
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-511-455-41

Query Match
Best Local Similarity 4.1%; Score 62.5; DB 6; Length 956;
Matches 47; Conservative 28; Mismatches 62; Indels 107; Gaps 10;

QY 26 RLPPFCRGMBEIA-----NFPVRPSDVWVITPKSGTSLLOEVVYLVVSQGADPDEIGLM 80
Db 443 RYEGFCVDMKELAEILRFNYKIRLVGDGVGYVPEANGTWTGNGELIARKADLAVAGLT 502
QY 81 NIDEQLPVLEYPOP-----GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYARNPK 136
Db 503 ITAREKVIDFSKPFMTLGISIL-----YR-----IHMGRKP- 534
QY 137 DLVSVYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVFQEFWEHRMDSNVLFKY 196
Db 535 ----GYFSF-----LDPFSPGW-----LFM-- 551
QY 197 EDMHRDLVTMVEQLARFLGVSCDKAQLAEALTEH-----CHQLVDQCCNAEAL- 243
Db 552 -----LLAYLAVSCVLFVARLTPYEWYSPHPCAQGRCNLLVNGYSLGSLW 598
QY 244 -PVG 246
Db 599 FPVG 602

RESULT 22
US-11-283-522-34
; Sequence 34, Application US/11283522
; Publication No. US2006008914A1
; GENERAL INFORMATION:
; APPLICANT: Pan, James
; APPLICANT: Adams, Sean
; TITLE OF INVENTION: UCP5
; FILE REFERENCE: P1663R2
; CURRENT APPLICATION NUMBER: US/11/283,522
; PRIOR FILING DATE: 2005-11-18
; PRIOR APPLICATION NUMBER: US/10/270,861
; PRIOR FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US/09/433,622
; PRIOR FILING DATE: 1999-11-02
; PRIOR APPLICATION NUMBER: US 60/110,286
; PRIOR FILING DATE: 1998-11-30
; PRIOR APPLICATION NUMBER: US 60/129,583
; PRIOR FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: US 60/143,886
; PRIOR FILING DATE: 1999-07-15
; NUMBER OF SEQ ID NOS: 36
; SEQ ID NO 34
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Homo Sapien
US-11-283-522-34

Query Match
Best Local Similarity 4.1%; Score 62; DB 7; Length 309;
Matches 30; Conservative 18; Mismatches 43; Indels 28; Gaps 6;

QY 97 DIIKE-LTSPRLIKSHLPYRFLPSDLHNGDSKVIYARNPKDLVVSY------QF---- 145
Db 198 DLIKDALIKANLMTDDLPCFTSA---FGAGFCTTIVIASPVDVVKTRYMNSALGOYSSAG 254
QY 146 HRSRLTMSYRGTFQFCRRFMNDKLGYSWFHVFQEFWEHRMDSNVLFKYEDMHRDLV 204
Db 255 HCAITMLQKEGP-RAFYKGFMPSLRLGWS-----NVVMFVTEYQLKRALM 299

RESULT 23
US-10-505-928-642
; Sequence 642, Application US/10505928
```


QY 69 SQGADPDEIGLMNIDQLPVLVYPPQGLDIIKELTSPR--LIKSHLPYRFLPSD--LHNG 124
Db 124 -----KLRGNFPFSDGVMQKKTGWEASSRMYPEDGALKGEMKMLKLDGGHYDA 176
QY 125 DSKVIYMARNPDKDLVVS 142
Db 177 EVKTYMAKPVQLPGAY 194

RESULT 27

US-10-505-928-233
; Sequence 233, Application US/10505928
; Publication No. US2006008832A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 233
; LENGTH: 512
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-233

Query Match 4.0%; Score 61.5; DB 6; Length 512;
Best Local Similarity 22.1%; Pred. No. 41;
Matches 29; Conservative 19; Mismatches 36; Indels 47; Gaps 6;
QY 62 QBWVYLVSGQADPDE-IGLMNIDEOLPVLEY-PQGLDIIKELTSPRLIKSHLPYRFLPS 119
Db 212 QELLSLVNINNFGVVGSGNPADIPILRYLPNFSLNAFKOLNE----- 256
QY 120 DLHNGDSKVIYMARNPDKDLVVSYYOYFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFH 179
Db 257 -----KFYSFMQKVVKEHYK-IFEKGHIRDITDSL-----IEHC 289
QY 180 QFWEHRMDSN 190
Db 290 Q--EKOLDEN 297

RESULT 28

US-10-511-937-2512
; Sequence 2512, Application US/10511937
; Publication No. US2006008836A1
; GENERAL INFORMATION:
; APPLICANT: EXPRESSION DIAGNOSTICS, INC.
; APPLICANT: Wohlgenuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; APPLICANT: Prentice, James
; APPLICANT: Morris, MacDonald
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; FILE REFERENCE: 50661200104
; CURRENT APPLICATION NUMBER: US/10/511,937
; CURRENT FILING DATE: 2004-10-19
; PRIOR APPLICATION NUMBER: PCT/US2003/012946
; PRIOR FILING DATE: 2003-04-24
; PRIOR APPLICATION NUMBER: US 10/131,831
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US 10/325,899
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 3117
; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 2512
; LENGTH: 789
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-511-937-2512

Query Match 4.0%; Score 61.5; DB 6; Length 789;
Best Local Similarity 22.5%; Pred. No. 79;
Matches 32; Conservative 23; Mismatches 38; Indels 49; Gaps 8;
QY 30 FCRGKMEETIANFPVRPSDVWIVTYPKSGTSLQLQEVVYLVVSQAGADPDEIGLMNIDQLPVL 89
Db 164 YCRDPAERL-HYPY-PGELTMNLTQTOSLQK-----PSTEKNEICPKNV-----PKR 210
QY 90 EY-----POPGLDI-----IKELTS-----PRL-----IKSHLP 113
Db 211 EYSVKEILKDSNPSKGLYRSNISPLTSEKLDLDFRRRSGPEMPFYPRVVYPPIRAPLP 270
QY 114 YRFLPSDLHNGDSKVIYMARNP 135
Db 271 EDFLKASLAYGIERPTYITRSP 292

RESULT 29

US-11-211-917-40
; Sequence 40, Application US/11211917
; Publication No. US20060093600A1
; GENERAL INFORMATION:
; APPLICANT: BEDIAN, VAHE
; APPLICANT: GLADUE, RONALD P.
; APPLICANT: CORVALAN, JOSE
; APPLICANT: JIA, XIAO-CHI
; APPLICANT: FENG, XIAO
; TITLE OF INVENTION: ANTIBODIES TO CD40
; FILE REFERENCE: ABX-PF/3 US
; CURRENT APPLICATION NUMBER: US/11/211,917
; PRIOR FILING DATE: 2005-08-25
; PRIOR APPLICATION NUMBER: US/10/292,088
; PRIOR FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: 60/348,980
; NUMBER OF SEQ ID NOS: 147
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 40
; LENGTH: 239
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-211-917-40

Query Match 4.0%; Score 61; DB 7; Length 239;
Best Local Similarity 22.8%; Pred. No. 14;
Matches 39; Conservative 25; Mismatches 57; Indels 50; Gaps 11;
QY 1 MAESEATPSTPGEFESKYFEHGVRLPPFCRGKMEETIANFPVRPSDVWIVTYPKSGTSL 60
Db 24 MTQSPSLSPVTPGEPAS-----ISCRSSQSVLYSNGYNYLD-WYLKQPGSQPL 71
QY 61 LQEVVYLVVSQAD--PDE-----IGLMNIDEQ-----LPVLEYP---QPG-- 95
Db 72 L---IYLGSRASGVDFRFGSGSGTDTLTKISRVAEDVGYYCMQVLQTFPTFGPGTK 128
QY 96 LDIKELTSPRLIKSHLPYRFLPSD--LHNGDSKVIYMARNP--PKDLVVS 142
Db 129 VDKRTVAAPSV-----FIFFPSDEQLKSGTASVVCLLNNFYPREAKVQW 173

RESULT 30

US-10-511-937-2584
; Sequence 2584, Application US/10511937
; Publication No. US2006008836A1
; GENERAL INFORMATION:
; APPLICANT: EXPRESSION DIAGNOSTICS, INC.
; APPLICANT: Wohlgenuth, Jay

Publication No. US20060088906A1
; GENERAL INFORMATION:
; APPLICANT: Neose Technologies, Inc.
; APPLICANT: DeFrees, Shawn
; APPLICANT: Zopf, David
; APPLICANT: Bayer, Robert
; APPLICANT: Hakes, David
; APPLICANT: Chen, Xi
; APPLICANT: Bowe, Caryne
; TITLE OF INVENTION: ERYTHROPOIETIN: REMODELING AND
; FILE REFERENCE: 040853-01-5083-US02
; CURRENT APPLICATION NUMBER: US/11/183,218
; CURRENT FILING DATE: 2005-07-15
; PRIOR APPLICATION NUMBER: US 10/410,945
; PRIOR FILING DATE: 2003-04-09
; PRIOR APPLICATION NUMBER: PCT/US02/32263
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: US 60/407,527
; PRIOR FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: US 60/404,249
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 60/396,594
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: US 60/391,777
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/387,292
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: US 60/344,692
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: US 60/334,301
; PRIOR FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: US 60/334,233
; PRIOR FILING DATE: 2001-11-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 30
; LENGTH: 2351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-183-218-30

Query Match 3.9%; Score 60; DB 7; Length 2351;
Best Local Similarity 20.0%; Pred. No. 6.1e+02;
Matches 51; Conservative 39; Mismatches 85; Indels 80; Gaps 14;
QY 53 YPKSGTSLQEVVYLVSGADPDEIGLMNIDEQLPVLEYQPGLDIIKELTSPLRIK--S 110
Db 473 YGEVGDITLL---IIFKNQASRPYNI-----YPH-GITDVRPLYSRRLPKGVK 515
QY 111 HL-PYRFLPSDL-----HNGDSKVIYMARNPKDLVSVYQFHRSLRTWS----- 153
Db 516 HLKDPFLPGEIFKYKWTVTVEDGPTK-----SDPCLTRYYSFVNMRDLASGLIGPL 570
QY 154 ---YRGTFQFCRRFMNDKLGXGWPHEVQFWEHRMDSNVLFLKYEDMHRDLVTMVEQL 210
Db 571 LICYKESVDQRGNQIMSDK-----RNVILFSVFENRSWY--LTENI 610
QY 211 ARFL-----GVSCDKAQLEALTEHCHQLVDQCNAEALPVGRGVGLWKDIFTVSMNEKFD 266
Db 611 QRFLPNPAGVQLEDPEFOA-SNIMHSINGYVFDLSQLSVCLHEVAYW---YILSIGAQT 666
QY 267 LV-----YKQKM 273
Db 667 FLVSFFSGYTFKHKM 681

RESULT 34
US-11-280-757-35
; Sequence 35, Application US/11280757
; Publication No. US20060099685A1
; GENERAL INFORMATION:

; APPLICANT: Crucell Holland B.V.
; APPLICANT: Bout, Abraham
; APPLICANT: Opstelten, Dirk-Jan
; APPLICANT: Yallop, Christopher
; TITLE OF INVENTION: RECOMBINANT EXPRESSION OF FACTOR VIII IN HUMAN CELLS
; FILE REFERENCE: 0034 D US P00 CIP
; CURRENT APPLICATION NUMBER: US/11/280,757
; CURRENT FILING DATE: 2005-11-15
; PRIOR APPLICATION NUMBER: US 10/234,007
; PRIOR FILING DATE: 2002-09-03
; PRIOR APPLICATION NUMBER: US 09/549,463
; PRIOR FILING DATE: 2000-04-15
; PRIOR APPLICATION NUMBER: 60/129,452
; PRIOR FILING DATE: 1999-04-15
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 35
; LENGTH: 2351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-280-757-35
Query Match 3.9%; Score 60; DB 7; Length 2351;
Best Local Similarity 20.0%; Pred. No. 6.1e+02;
Matches 51; Conservative 39; Mismatches 85; Indels 80; Gaps 14;
QY 53 YPKSGTSLQEVVYLVSGADPDEIGLMNIDEQLPVLEYQPGLDIIKELTSPLRIK--S 110
Db 473 YGEVGDITLL---IIFKNQASRPYNI-----YPH-GITDVRPLYSRRLPKGVK 515
QY 111 HL-PYRFLPSDL-----HNGDSKVIYMARNPKDLVSVYQFHRSLRTWS----- 153
Db 516 HLKDPFLPGEIFKYKWTVTVEDGPTK-----SDPCLTRYYSFVNMRDLASGLIGPL 570
QY 154 ---YRGTFQFCRRFMNDKLGXGWPHEVQFWEHRMDSNVLFLKYEDMHRDLVTMVEQL 210
Db 571 LICYKESVDQRGNQIMSDK-----RNVILFSVFENRSWY--LTENI 610
QY 211 ARFL-----GVSCDKAQLEALTEHCHQLVDQCNAEALPVGRGVGLWKDIFTVSMNEKFD 266
Db 611 QRFLPNPAGVQLEDPEFOA-SNIMHSINGYVFDLSQLSVCLHEVAYW---YILSIGAQT 666
QY 267 LV-----YKQKM 273
Db 667 FLVSFFSGYTFKHKM 681
RESULT 35
US-11-101-316-140
; Sequence 140, Application US/11101316
; Publication No. US20060099657A1
; GENERAL INFORMATION:
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: ANTIBODIES TO A POLYPEPTIDE ENCODED BY A NUCLEIC ACID
; FILE REFERENCE: P3230R1C17C1
; CURRENT APPLICATION NUMBER: US/11/101,316
; CURRENT FILING DATE: 2005-04-06
; PRIOR APPLICATION NUMBER: 10/063526
; PRIOR FILING DATE: 2002-05-03
; PRIOR APPLICATION NUMBER: 10/006867
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 09/380137
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: PCT/US99/12252
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: 60/087759

```
; PRIOR FILING DATE: 1998-06-02
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 140
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo Sapien
US-11-101-316-140

Query Match      3.9%; Score 59.5; DB 7; Length 310;
Best Local Similarity 20.5%; Pred. No. 32;
Matches 46; Conservative 26; Mismatches 77; Indels 75; Gaps 11;

Qy  2 AESRAETPS--TPGESKSYFHFHGVRLPPFCRGKMEIEANFPVRPSDVWIVTPYKSGTSL 60
Db  38 AEHGAEPAPPSAGASS-----NWTLPPLPFSKVVILIDALRDFV-----FGSKGVKF 89

Qy  61 LQEVVYLSQGAADPEIGLMNIDEQLPVLEYPO-----PG-LDIKELTSPRLIKS 110
Db  90 MPYTTLVEKGASHSFVA-----EAKPPTVTMPRIKALMTGSLPGFVDVIRNLNSPALL-- 143

Qy  111 HLPYRFLPSDLHNGDSKVIYARNPKOLVVSY-----YQFHSRLTMSYRGTFQEF- 161
Db  144 -----EDSVIROAKAAGKRVIFYGDETWVKLFPKH-----FVEYDGTTSFFV 185

Qy  162 -----CRRFNWDKLYGSGWFHGVQEFWEHRMDSNVLFKY 196
Db  186 SDYTEVDNNVTRHLDKVLKRGDW-----DILILHY 215
```

```
RESULT 36
US-10-196-749-404
; Sequence 404, Application US/10196749
; Publication No. US20060094864A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C340
; CURRENT APPLICATION NUMBER: US/10/196,749
; CURRENT FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 404
```

```
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-196-749-404

Query Match      3.9%; Score 59.5; DB 6; Length 347;
Best Local Similarity 31.4%; Pred. No. 37;
Matches 16; Conservative 10; Mismatches 12; Indels 13; Gaps 4;

Qy  27 LPPF-CRGKMEIEANFPVRPSDVWIV--TYPKSGTSLQEVVYLSQGADP 74
Db  295 IPPPQCRHRCQSA-MBIEPGDIGYDVTTHWK-----VYVIARGVQP 335

RESULT 37
US-11-101-316-124
; Sequence 124, Application US/11101316
; Publication No. US20060099657A1
; GENERAL INFORMATION:
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: ANTIBODIES TO A POLYPEPTIDE ENCODED BY A NUCLEIC ACID
; FILE REFERENCE: P3230R1C17C1
; CURRENT APPLICATION NUMBER: US/11/101,316
; CURRENT FILING DATE: 2005-04-06
; PRIOR APPLICATION NUMBER: 10/063526
; PRIOR FILING DATE: 2002-05-03
; PRIOR APPLICATION NUMBER: 10/006867
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 09/380137
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: PCT/US99/12252
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 124
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Homo Sapien
US-11-101-316-124

Query Match      3.9%; Score 59.5; DB 7; Length 513;
Best Local Similarity 24.4%; Pred. No. 68;
Matches 44; Conservative 19; Mismatches 50; Indels 67; Gaps 11;

Qy  31 CRGKMEIEANFPVRPSDVWIVTPYKSGTSLQEVVYLSQGAADPEIGLMNIDQLPVLE 90
Db  40 CEGKMK-----VYCES--QKLQEIPTSSISAGC-----LGLSLRYNSLQKLK 77

Qy  91 YPO-PGLDIIKELTSPRLIKSHLPYRFLPSDLHNG-----DSKVIYMARNP----- 135
Db  78 YNQPKGLN---OLTWLYLDHNHISN--IDENAFNGIRRLKELIILSSNRISFLNNTFRPV 132

Qy  136 ---KDLVVSYQYFH-----RSLRTMSYRGTFQEFRRFMNDKLG 172
Db  133 TNLRLNLSYNQLHSLGSEQFRGLRKULSLHRSNLSRTIPVR-IFQD-CRNLELLDLY 190

RESULT 38
US-11-252-080-2
; Sequence 2, Application US/11252080
; Publication No. US20060090217A1
; GENERAL INFORMATION:
; APPLICANT: Texas A&M University
; TITLE OF INVENTION: ISOLATION OF SUGARCANE PROTEINS INVOLVED IN POSTTRANSCRIPTIONAL
; TITLE OF INVENTION: SILENCING AND METHODS OF USE
```



```
; LENGTH: 1005
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-473-691B-2

Query Match      3.8%; Score 58; DB 6; Length 1005;
Best Local Similarity 28.6%; Pred. No. 2.8e+02;
Matches 16; Conservative 11; Mismatches 21; Indels 8; Gaps 2;

QY 42 PVRSDVWIVTPKSGTSLQEVVYVLSQADPDEIGLNMNIDEQLPVLEYPQGLD 97
||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 661 PVEIQNPYLHVEPE-----KEWSLVPELSPEVDG--NASTENSMPAPSPGLD 708

RESULT 51
US-10-505-928-662
; Sequence 662, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; PRIOR FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 662
; LENGTH: 1504
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-662

Query Match      3.8%; Score 58; DB 6; Length 1504;
Best Local Similarity 20.3%; Pred. No. 5.1e+02;
Matches 55; Conservative 39; Mismatches 79; Indels 98; Gaps 17;

QY 82 IDEQLPV-----LEYQPGLDIKELTSRLIKSHLPYRFLPSDLHNGD---SK 127
||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1086 VPEETPVEAPAPPAEAPQYPSHL--IQATSEENQIPSHLP--ACPSLRHVASLRGSA 1141

QY 128 VIYARNP-----KDLVSVYVQFHR-----LRTMSYRGTFQFCRFRMND 168
||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1142 IIEFLHSSIAEVENBELRLHMSVVFYQTPGLVTA CVLLSTKAVFYVLHDLRLRYFSE 1201

QY 169 KLGYSWFHFVQEFWEHRMD-----SNVLEFLKYEDMHRDLVTMVEQLARFLG--- 215
||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1202 PL-----QDFWHQKVTYNNSPFHISQCFLVLSLQSVNVLGDFDQHPRLTGSTPM 1252

QY 216 --VSCDKAQLEALTEHC---HQLV-----DQCNAEALPVGRVGLWKDIF-----TV 259
||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1253 QVVTC--LTRDSYLTHCFLOHLMVVLSSLERTPSP--PVD-----KDFYSERGNKTT 1301

QY 260 SNNEKFDLVY-----KQMGKCDLTF 280
||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1302 GKMYNELIHSRVRKFTYPSSEIEG--DLTF 1330

RESULT 52
US-11-251-466-2
; Sequence 2, Application US/11251466
; Publication No. US20060094090A1
; GENERAL INFORMATION:
; APPLICANT: E.I. duPont de Nemours and Company
; APPLICANT: Xue, Zhixiong
; APPLICANT: Damude, Howard Glenn
; TITLE OF INVENTION: A MORTIERELLA ALPINA LYSOPHOSPHATIDIC ACID ACYLTRANSFERASE
; TITLE OF INVENTION: HOMOLOG FOR ALTERATION OF POLYUNSATURATED FATTY ACIDS AND OIL
; TITLE OF INVENTION: CONTENT IN OLEAGINOUS ORGANISMS
; FILE REFERENCE: CL2793 USNA
; CURRENT APPLICATION NUMBER: US/11/251,466
; CURRENT FILING DATE: 2005-10-14
```

```
; PRIOR APPLICATION NUMBER: US 60/624812
; PRIOR FILING DATE: 2004-11-04
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2
; LENGTH: 308
; TYPE: PRT
; ORGANISM: Mortierella alpina
US-11-251-466-2

Query Match      3.8%; Score 57.5; DB 7; Length 308;
Best Local Similarity 19.5%; Pred. No. 52;
Matches 31; Conservative 28; Mismatches 61; Indels 39; Gaps 9;

QY 140 VSYVQFHRSLRTMSYRGTFQFCRFRMNDKL-----GYGSWFHFVQEFWEHRMDSNVLF 193
||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 93 VDFYMFH-----TVAARRGMLNYMKYFAKDSLKYIPFYGWMWI-----MGMLF 136

QY 194 LKYEDMHRDLVTMVEQLARFLGVSCDKAQLEALTEHCHQLVDQCCNAEALPVGRVGLW 253
||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 137 IN-RNWQDQKINKMFARILIDIA-PWVASFLEGSRLTSPKLAASQKFMLG-LPLL 193

QY 254 KDI-----FTVSMNEKFDLVYKQMGKC--DLTFDFY 283
||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 194 SNVAMPRTKGFACVKN-----FRGTHVKCVYDFTFAYY 227

RESULT 53
US-10-505-928-646
; Sequence 646, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 646
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-646

Query Match      3.8%; Score 57.5; DB 6; Length 609;
Best Local Similarity 20.9%; Pred. No. 1.5e+02;
Matches 36; Conservative 30; Mismatches 57; Indels 49; Gaps 10;

QY 107 LIKSHL-----PYRFLPSDLHNGDSKVIYARN-----PKDLVSVYVQFHR 147
||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 452 LTKDHI FLNSALRATAPYKFL-SEQHG--VKAVFMNKHWDYEEYPTTSSFVLNGFMV-- 506

QY 148 SLRTMSYRGTFQFCRFRMNDKLGYSWFHFVQEFWEHRMDSNVLFKYEDMHRDLVTMV 207
||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 507 -----SLIGLYD--LKETAGEKLG-----KEARSLYERGMESLKAMLPDYDTGSG--TIY 552

QY 208 EQLARFLGVSCDKAQLEALTEHCHQL-----VDQCCNAEALPVGRVGLWK 254
||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 553 DLRHFMGLGAPNARWDYHTTHINQLQLLSTIDES-----PIKFEFVRWK 598

RESULT 54
US-10-505-928-416
; Sequence 416, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
```

```
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 416
; LENGTH: 708
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-416

Query Match      3.8%; Score 57.5; DB 6; Length 708;
Best Local Similarity 21.9%; Pred. No. 1.8e+02;
Matches 57; Conservative 23; Mismatches 89; Indels 91; Gaps 12;

Qy 4 SEATPTSPGPFESKY-----FEFHGVRLLPPFCRGKWEI--ANFVRPSDVWIV 51
Db 268 SSVVTSLSGPAVKKHVGLLRIGKRMNMHKIPLHTVRQFFMEDIVLANHP-----DIFNP 323
Qy 52 TYPKSGTSLQEVVYLVVLSQGDADPEIGLNMNIDEQLPVLEYQPQGLDIIKELTSRLIKSH 111
Db 324 DNPKV-TQAISFC-----LEKIEEMLENAE-----RERLGNSH 356
Qy 112 LPYR-----FLPSDLHNGDSKVIYIYMARNPKDLVSVYQFHR----- 147
Db 357 QPEKPLVRLRVVDYSGGFPPFVLRFSQKFDVRVANPKDII--HPRHREQKEKTGEENF 414
Qy 148 -SLRTMSYRGT-----FQFCRRFNMNDKLGYSWFEHVOEFWE-HRMDNSNVLF 194
Db 415 GKLIKPSGEGTLRVDELVKVQYFQAENQVLSLLTERGMGEAVQEFVDKKEKDAIELV 474
Qy 195 KYEDMHRDLVTMVEQLARFL 214
Db 475 KYQ-----LEKTQRF 485

RESULT 55
US-10-504-120-21
; Sequence 21, Application US/10504120
; Publication No. US2006008829A1
; GENERAL INFORMATION:
; APPLICANT: Exelixis, Inc.
; TITLE OF INVENTION: MINRs AS MODIFIERS OF INSULIN RECEPTOR SIGNALING AND METHODS OF
; TITLE OF INVENTION: USE
; FILE REFERENCE: EX03-003C-PC
; CURRENT APPLICATION NUMBER: US/10/504,120
; CURRENT FILING DATE: 2004-08-06
; PRIOR APPLICATION NUMBER: 60/354,824
; PRIOR FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: 60/358,217
; PRIOR FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: 60/358,189
; PRIOR FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: 60/358,126
; PRIOR FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: 60/358,995
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 60/358,756
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 60/358,765
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 60/359,531
; PRIOR FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 60/360,222
; PRIOR FILING DATE: 2002-02-26
; PRIOR APPLICATION NUMBER: 60/360,224
; PRIOR FILING DATE: 2002-02-26
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 21
; LENGTH: 1763
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-504-120-21

Query Match      3.8%; Score 57.5; DB 6; Length 1763;
Best Local Similarity 20.1%; Pred. No. 7.4e+02;
Matches 54; Conservative 41; Mismatches 100; Indels 73; Gaps 13;

Qy 36 EEIANFVRPSDVW-----IVTYPKSGTSLQEVV-----YLVSQGDADPD-EI 77
Db 158 EELADFVLQWMDVGLSSEFLLVLNLFKFSYDEVIARMVQMICLLCVRTASSVDIEV 217
Qy 78 GLMNID-----EQLPVLEYQPQGLDIIKELTS-----RLIKSHLPYRFLPSDL 121
Db 218 SLQVLDVAVCYNCILPASLPLFIVTLCTINVKELCPCKWKLMMNLLGTHLGHSAIYNMC 277
Qy 122 HNGDSKVIYMARNP--KDLV-----VSYQFHR--SLRTMSYRGTFQFCRRFM--NOKLG 171
Db 278 HLMDRA-YMEDAPLLRGAVFFVGMLWGAHRLYSLRN-SPTSVFSPFSYQAMACPNVVS 335
Qy 172 YGSWFEHVQEFWEHRMDSNVLF-----KVEDMHRDLVTMVEOLA 211
Db 336 YEIVLSITRLIKYRKELQVAVMDILLNIIRLLOQLDLSPELTIIVHDLTTVEEL- 394
Qy 212 RFLGVSCDKAQLEALTEHCHQLVDCCN 239
Db 395 -----CDQNEFHGSGQERYFELVERCAD 416

RESULT 56
US-10-505-928-288
; Sequence 288, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 288
; LENGTH: 720
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-288

Query Match      3.7%; Score 57; DB 6; Length 720;
Best Local Similarity 20.3%; Pred. No. 2.2e+02;
Matches 56; Conservative 36; Mismatches 94; Indels 90; Gaps 14;

Qy 32 RGMKEETANFVRPSDVWIVTYPKSGTSLQEVVYLVVLSQGDADPEIGLNMNIDEQLPVL-- 89
Db 265 KASMENQPHSKPLPSDIAVIMYTSGSTGLPKGVMIHSNII-----AGITGMAERIPELGE 320
Qy 90 EYPQPG---LDIIKELTSPLIKSH---LPYRFLPSDLHNGDSKV----- 128
Db 321 EDVYIGYLPPLAHVLELSAELVCLSHGCRIGYS-SPQTLADQSSKIKGSKGDSMLKPTL 379
Qy 129 -----IYMARNPKDLVSVYQFHRSLRTMSY-----RGTFQFCRRFMNDKL 170
Db 380 MAAVPEIMDRIV--KNVWNKYSEMSSFQNLFILAYNYKMEQISKGRNTPLCDSFV---- 433
Qy 171 GYGSWFEHVQEFWEHRMDSNVLFKYEIDMHRDLVTMVEQLARFLGVSCDKAQLEALTEHC 230
Db 434 -----FRKVRSL-----LGGNIRLL-----LCGGAPLSATTORF 462
Qy 231 HQLVDQCCNAEALPVGRGVRGLMKWDIFTVSNNEKFD 266
Db 463 MNIC-FCC-----PVQOG-YGLTESAGAGTISEVWD 491
```

```
RESULT 57
US-10-505-928-491
; Sequence 491, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 491
; LENGTH: 832
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-491

Query Match      3.7%; Score 57; DB 6; Length 832;
Best Local Similarity 20.9%; Pred. No. 2.7e+02;
Matches 38; Conservative 29; Mismatches 57; Indels 58; Gaps 8;

QY 100 KELTSR-RLI-----KSHLPYRFLPSDLHGDSKVIYMARNPKDLVSVYYQFHRSLRTWS 153
Db 109 KEITLPSRLIYYINODSESPYHVL-----DTKARHQKHNAKAVHLAQASQF----- 154
QY 154 YRGTFQECRRFRFMDNDKLGYSWFHVFQEFWEHRMDSNVFLKYEDMHRDLVTMVQELARF 213
Db 155 -----TEAFQSKFILDLI-----LNGLLSSDYVEIHYE----- 183
QY 214 LGVSCDQKALEALTEHC--HQLVDQCNAE--ALPVGGRGVGLWKDIFTVSMNEKFDLVYK 270
Db 184 -----NGKPYSGGEGHYHGIRGVKDSKVALSCNGLHGMFEDDTFYVMIEPLELVHD 239
QY 271 QK 272
Db 240 EK 241

RESULT 58
US-10-505-928-35
; Sequence 35, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 35
; LENGTH: 926
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-35

Query Match      3.7%; Score 57; DB 6; Length 926;
Best Local Similarity 28.4%; Pred. No. 3.2e+02;
Matches 21; Conservative 16; Mismatches 33; Indels 4; Gaps 3;

QY 5 EAETSTP--GFESKYFEFHGVRLLPPFCRGKMBEIANFPVR-PSDVWLVTVYKSGTSL 61
Db 249 KTEITLSPKDSBKSKKGLKFLRL--SPKKDKTKQLANFSAQFPPEWPLRDEDTPATIP 307
QY 62 QEVVYLVSGADPD 75
Db 308 REVEWEIIRINPD 321

RESULT 59
US-10-505-928-100
; Sequence 100, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 100
; LENGTH: 2871
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-100

Query Match      3.7%; Score 57; DB 6; Length 2871;
Best Local Similarity 20.4%; Pred. No. 1.7e+03;
Matches 47; Conservative 40; Mismatches 91; Indels 52; Gaps 11;

QY 1 MAESAETPSTPGEPESKYFEFH--GVRLLPPFCRGK-----MBEIANFPVRPSDVWIVT 52
Db 2222 LVDSGILRFPSTVNELESQISYDEVGERIKDFLQSSCIAGIYNETTKQKLGIVEAMKIG 2281
QY 53 YPKSGTSL-LQEVVYLVSGADPDDEIGLNMNIDEQLPVLEYPOPLDII-----KELTSRPL 107
Db 2282 LVRPGTALALELQAQAATGFIVDP-----VSNL--RLPVEEAYKRGVLGVEFKELLSAERA 2335
QY 108 IKSHLPYRFLPSDLHGDSKVIYMARNPKDL-----VSVYYQFHRSL 149
Db 2336 VTGY-----NDPETHNIISLFQAMN--KELIEKHGIRLLEAQATGDIIDPKSHRLP 2387
QY 150 RTMSY-RGTQFQECRRFRFMDNDKLGYSWFHVFQEFWEHRMDSNVFLKYED 198
Db 2388 VDIAYKRGVFNELSELSDP-----SDTKGFFDPNTEENTLYQLKE 2431

RESULT 60
US-10-505-928-167
; Sequence 167, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 167
; LENGTH: 493
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-167

Query Match      3.7%; Score 56.5; DB 6; Length 493;
Best Local Similarity 18.7%; Pred. No. 1.4e+02;
Matches 63; Conservative 43; Mismatches 108; Indels 123; Gaps 17;

QY 19 YFEFHGVRLLPPFCRGKMBEIANFPVRPSDVWIVTVYKSGTSLLOEVVYLVSGADPD 78
Db 161 YLSFHKLL--HLQGERE-----PG--WI---KQLFTNFISFTLKLVLKGQICKIN 205
QY 79 LM-NI-----DEQLPVLEYQPGLDIIKELTSRLI-----KSHLPYRFLPSDLH 122
Db 206 VISNIMADFVQTRAASILSDGIDGVDI--SLTGDPFVITASYLESHHKGHFIYKNVSEDL 263
```


[illegible]

RESULT 61

```

US-10-511-937-2991
; Sequence 2991, Application US/10511937
; Publication No. US2006008836A1
; GENERAL INFORMATION:
; APPLICANT: EXPRESSION DIAGNOSTICS, INC.
; APPLICANT: Wohlgemuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; APPLICANT: Prentice, James
; APPLICANT: Morris, MacDonald
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND MONITORING TRANSPLANT REJECTION
; FILE REFERENCE: 506612000104
; CURRENT APPLICATION NUMBER: US/10/511,937
; CURRENT FILING DATE: 2004-10-19
; PRIOR APPLICATION NUMBER: PCT/US2003/012946
; PRIOR FILING DATE: 2003-04-24
; PRIOR APPLICATION NUMBER: US 10/131,831
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US 10/325,899
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 3117
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2991
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-511-937-2991

```

```

Query Match      3.7%; Score 56.5; DB 6; Length 505;
Best Local Similarity 29.3%; Pred. NO. 1.4e+02;
Matches 29; Conservative 9; Mismatches 22; Indels 39; Gaps 7;

Qy 37 EIANPPVR-----PSDVW-----IVTY----PKSGTSLLEGVVLVSQ 71
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 394 EGAKFPKWTAEAINFGSFTIKSDVMSFGILLMEIVTYGRIPYPMGS-NPEVIRALERG 452
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

Qy 72 AD-----PDEI-GLM-----NIDEQLPVLEYQPGLD 97
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 453 YMRPRPENCPELYNIIMMRKWKNRPEERPTFEYIQSVLD 491
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

```

RESULT 62

```

US-11-301-924-4
; Sequence 4, Application US/11301924
; Publication NO. US20060090218A1
; GENERAL INFORMATION:
; APPLICANT: Joshua Z. Levin
; APPLICANT: Ken Phillips
; APPLICANT: Greg Budziszewski
; APPLICANT: Fred Meins
; APPLICANT: Zhenya Glazov
;
Best Local Similarity 31.4%; Pred. No. 24;
Matches 22; Conservative 16; Mismatches 30; Indels 2; Gaps 2

```

Db 124 Q-HRISATSS 132

RESULT 64
US-11-264-784-54
; Sequence 54, Application US/11264784
; Publication No. US20060094092A1
; GENERAL INFORMATION:
; APPLICANT: E.I. duPont de Nemours & Co., Inc.
; APPLICANT: Damude, Howard Glenn
; APPLICANT: Gillies, Peter John
; APPLICANT: Macool, Daniel Joseph
; APPLICANT: Picataggio, Stephen K.
; APPLICANT: Pollak, Dana M. Walters
; APPLICANT: Ragnianti, James John
; APPLICANT: Xue, Zhixiong
; APPLICANT: Yadav, Narendra S.
; APPLICANT: Zhang, Hongxiang
; APPLICANT: Zhu, Quinn
; TITLE OF INVENTION: HIGH ARACHIDONIC ACID PRODUCING STRAINS OF YARROWIA LIPOLYTICA
; FILE REFERENCE: CL3136 USA
; CURRENT APPLICATION NUMBER: US/11/264,784
; CURRENT FILING DATE: 2005-11-01
; NUMBER OF SEQ ID NOS: 375
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 54
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Mortierella alpina
US-11-264-784-54

Query Match 3.7%; Score 56; DB 7; Length 275;
Best Local Similarity 21.7%; Pred. No. 64;
Matches 13; Conservative 16; Mismatches 27; Indels 4; Gaps 2;

Qy 139 VVSYQFHRSLTMSYRGTFQFEC---RRFMNDKLGYSWFEHVOEFWEHRMDSNVLFK 195
Db 85 IFYMFPAWVKNFRTHLHEAYCDTQSLWNNALGYWGLFYLSKPYE-VIDTIIILK 143

RESULT 65
US-10-489-730-4
; Sequence 4, Application US/10489730
; Publication No. US20060088825A1
; GENERAL INFORMATION:
; APPLICANT: Melino, Gennaro
; APPLICANT: Hayes, Ian
; APPLICANT: de Laurenzi, Vincenzo
; APPLICANT: Barcaroli, Daniela
; APPLICANT: Candi, Eleonora
; APPLICANT: Bernassola, Francesca
; APPLICANT: Tobler, Andreas
; APPLICANT: Novak, Urban
; TITLE OF INVENTION: Human Delta-N p73 Molecules and Uses Thereof
; CURRENT APPLICATION NUMBER: US/10/489,730
; CURRENT FILING DATE: 2004-03-16
; PRIOR APPLICATION NUMBER: PCT/GB02/04238
; PRIOR FILING DATE: 2002-09-17
; PRIOR APPLICATION NUMBER: US 60/322,436
; PRIOR FILING DATE: 2001-09-17
; NUMBER OF SEQ ID NOS: 39
; SEQ ID NO 4
; LENGTH: 450
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: peptide sequence of delta-N p73-beta protein
US-10-489-730-4

Query Match 3.7%; Score 56; DB 6; Length 450;
Best Local Similarity 29.7%; Pred. No. 1.4e+02;
Matches 22; Conservative 15; Mismatches 29; Indels 8; Gaps 4;

Qy 62 QEVVYLVSQGADPDDEIGLMNIDEQLPVL E-YPOQGLDIIKELTSRLIKSHL-PYRFLP- 118
Db 303 EDTYYLVQVRGRENFEI-LMKLKESLELMELVPQLVDSYRQQQLLQRPSPSHLPQPSYGVF 361

Qy 119 ---SDLHNGDSKV 128
Db 362 LSPMKVHGGMNKL 375

RESULT 66
US-10-528-032-17
; Sequence 17, Application US/10528032
; Publication No. US20060101530A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: RORS AS MODIFIERS OF THE p21 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX03-068C-US
; CURRENT APPLICATION NUMBER: US/10/528,032
; CURRENT FILING DATE: 2005-03-16
; PRIOR APPLICATION NUMBER: US 60/411,010
; PRIOR FILING DATE: 2002-09-16
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 17
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-528-032-17

Query Match 3.7%; Score 56; DB 6; Length 459;
Best Local Similarity 22.3%; Pred. No. 1.4e+02;
Matches 54; Conservative 37; Mismatches 85; Indels 66; Gaps 16;

Qy 56 SGTSLQVWYLVSQGADPDDEIGLMNIDEQLPVL EYPOQGLDI--IKEL-TSPRLIKSHL 112
Db 138 SGTANGHVIDL-----PKSEGYNVDSGQSPD--OSGLDMTGIKIQEPIYDLTSV 189

Qy 113 PYRFLPSDLHNGD-----SKVIYMARPNKDLVVSYQFHRSLTMSYRGTFQFECRRF 165
Db 190 PNLFTYSSFNNGQLAPGITMTTEIDRIQN---IISKH-----LTCQY--TMEE----- 233

Qy 166 MNDKLGYSWFEHVOE-----FWHRMDSNVLFKLYEDMHRDLVTWVQLARF 213
Db 234 ----LHQLAWQTHYEEIKAYQSKSREALWQCCAQIQTHTAIQY-----VVFEAKRITGF 283

Qy 214 LGVSCDKAQLREALTEHCHQ--LVQOC-----CNAEALPVGR-GRVGLWKDIFTVSM-NEK 264
Db 284 MEL-CONDQILLKSGCLEVVLVRCRAFNPNNLTVLFEGKYGGQMFKALGSDDLVNEA 342

Qy 265 FD 266
Db 343 FD 344

RESULT 67
US-11-301-554-336
; Sequence 336, Application US/11301554
; Publication No. US20060088527A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Kalos, Michael D.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Durham, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria

;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
;; FILE REFERENCE: 210121.478C21
;; CURRENT APPLICATION NUMBER: US/11/301,554
;; CURRENT FILING DATE: 2005-12-13
;; PRIOR APPLICATION NUMBER: US 10/283,017
;; PRIOR FILING DATE: 2002-10-28
;; PRIOR APPLICATION NUMBER: US 10/113,872
;; PRIOR FILING DATE: 2002-03-28
;; PRIOR APPLICATION NUMBER: US 10/017,754
;; PRIOR FILING DATE: 2001-10-29
;; PRIOR APPLICATION NUMBER: US 09/902,941
;; PRIOR FILING DATE: 2001-07-10
;; PRIOR APPLICATION NUMBER: US 09/849,626
;; PRIOR FILING DATE: 2001-05-03
;; PRIOR APPLICATION NUMBER: US 09/736,457
;; PRIOR FILING DATE: 2000-12-13
;; PRIOR APPLICATION NUMBER: US 09/702,705
;; PRIOR FILING DATE: 2000-10-30
;; PRIOR APPLICATION NUMBER: US 09/677,419
;; PRIOR FILING DATE: 2000-10-06
;; PRIOR APPLICATION NUMBER: US 09/671,325
;; PRIOR FILING DATE: 2000-09-26
;; PRIOR APPLICATION NUMBER: US 09/658,824
;; PRIOR FILING DATE: 2000-09-08
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 2157
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 336
;; LENGTH: 480
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-11-301-554-336

Query Match 3.7%; Score 56; DB 7; Length 480;
Best Local Similarity 22.1%; Pred. No. 1.5e+02;
Matches 48; Conservative 29; Mismatches 80; Indels 60; Gaps 12;
QY 29 PFCR---GKMEIANFPVPDWMIVTPYKSGTSLQEVVLYVSGADPDPDGLMNDIEQ 85
DB 86 PCGSSLDGLLTHGHPFLVQPGVTLVEYNFYS-WNLIANVLYLES-----PAGVGFSYDDK 140
QY 86 LPV---LEVPOGDLIIKELTSPRLIKSHLPYRPLPSDLHN-----GDSKV-IYM----- 131
DB 141 FYATNDTEVASGNFALQDF-----FRLEPEYKNNKLFITGSEYAGIYIPTLAV 189
QY 132 -----ARNPKDLVV-----SYYPHRSRLTMSYRGTFQFCRRFPMNDKLGYSWFHVQ 180
DB 190 LVMQDPSMNLQGLAVGNGLSSYEQNDNSLVVFAY-----YHGLLGNRLW-----S 234
QY 181 EFWEHRMDSNVLFLKYEDMHRDLVTMVEQLARFLGVS 217
DB 235 SLQTHCCSQNKCNF-YDNKDLCEVTNLQEVARIQVNS 270

RESULT 68
US-09-949-925-82
;; Sequence 82, Application US/09949925
;; Publication No. US20060099575A9
;; GENERAL INFORMATION:
;; APPLICANT: Rosen et al.
;; TITLE OF INVENTION: 67 Human secreted proteins
;; FILE REFERENCE: P202392
;; CURRENT APPLICATION NUMBER: US/09/949,925
;; CURRENT FILING DATE: 2001-09-12
;; PRIOR APPLICATION NUMBER: US 60/232,150
;; PRIOR FILING DATE: 2000-12-09
;; PRIOR APPLICATION NUMBER: PCT/US99/01621
;; PRIOR FILING DATE: 1999-01-27
;; PRIOR APPLICATION NUMBER: US 60/073,160
;; PRIOR FILING DATE: 1998-01-30
;; PRIOR APPLICATION NUMBER: US 60/073,159
;; PRIOR FILING DATE: 1998-01-30

;; PRIOR APPLICATION NUMBER: US 60/073,165
;; PRIOR FILING DATE: 1998-01-30
;; PRIOR APPLICATION NUMBER: US 60/073,164
;; PRIOR FILING DATE: 1998-01-30
;; PRIOR APPLICATION NUMBER: US 60/073,167
;; PRIOR FILING DATE: 1998-01-30
;; PRIOR APPLICATION NUMBER: US 60/073,162
;; PRIOR FILING DATE: 1998-01-30
;; PRIOR APPLICATION NUMBER: US 60/073,161
;; PRIOR FILING DATE: 1998-01-30
;; PRIOR APPLICATION NUMBER: US 60/073,170
;; PRIOR FILING DATE: 1998-01-30
;; NUMBER OF SEQ ID NOS: 298
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 82
;; LENGTH: 578
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; NAME/KEY: SITE
;; LOCATION: (326)
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;; NAME/KEY: SITE
;; LOCATION: (342)
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;; NAME/KEY: SITE
;; LOCATION: (444)
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-949-925-82
Query Match 3.7%; Score 56; DB 1; Length 578;
Best Local Similarity 18.4%; Pred. No. 2e+02;
Matches 35; Conservative 35; Mismatches 74; Indels 46; Gaps 9;
QY 35 MEEIANFPVPDWMIVTPYKSGTSLQEVVLYVSGADPDP--EIGLMNIDQLPVLEYP 92
DB 361 VEKEGYPPLRP-----ELIESAMLYRATGCDPTLLELG-RDAVESIEKISKV 406
QY 93 QPGLDIIKELTSPRLIKSHLPYRPLPSDLHNGDSKVLYMARNPDKLVVSYVQPHRSRLTM 152
DB 407 ECGPATTKDRLDHKL-DNRMESFFLAETV-----KYLILLFDPXNFI-----HNN----- 450
QY 153 SYRGTFQFCRRFPMNDKLGYSWFHVQEFWEHRMDSNVLF-----KYEDMHRDL 203
DB 451 --GSTFDVAITPYGECILGAGGIYFNTE---AHPIDPAALHCCORLKEQWEVEDLMREF 505
QY 204 VTWVEQLARF 213
DB 506 YSLKRSRSKF 515

RESULT 69
US-10-489-730-2
;; Sequence 2, Application US/10489730
;; Publication No. US20060088825A1
;; GENERAL INFORMATION:
;; APPLICANT: Melino, Gennaro
;; APPLICANT: Hayes, Ian
;; APPLICANT: de Laurenzi, Vincenzo
;; APPLICANT: Barcaroli, Daniela
;; APPLICANT: Candi, Eleonora
;; APPLICANT: Bernassola, Francesca
;; APPLICANT: Tobler, Andreas
;; APPLICANT: Novak, Urban
;; TITLE OF INVENTION: Human Delta-N p73 Molecules and Uses Thereof
;; FILE REFERENCE: 19319.002
;; CURRENT APPLICATION NUMBER: US/10/489,730
;; CURRENT FILING DATE: 2004-03-16
;; PRIOR APPLICATION NUMBER: PCT/GB02/04238
;; PRIOR FILING DATE: 2002-09-17
;; PRIOR APPLICATION NUMBER: US 60/322,436
;; PRIOR FILING DATE: 2001-09-17
;; NUMBER OF SEQ ID NOS: 39


```
; APPLICANT: Durham, Margarita
; APPLICANT: Carter, Darick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C21
; CURRENT APPLICATION NUMBER: US/11/301,554
; CURRENT FILING DATE: 2005-12-13
; PRIOR APPLICATION NUMBER: US 10/283,017
; PRIOR FILING DATE: 2002-10-28
; PRIOR APPLICATION NUMBER: US 10/113,872
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 10/017,754
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: US 09/902,941
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 09/849,626
; PRIOR FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: US 09/736,457
; PRIOR FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: US 09/702,705
; PRIOR FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: US 09/677,419
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 09/671,325
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US 09/658,824
; PRIOR FILING DATE: 2000-09-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2157
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1903
; LENGTH: 2209
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-301-554-1903

Query Match      3.6%; Score 55.5; DB 7; Length 2209;
Best Local Similarity 20.7%; Pred. No. 1.7e+03;
Matches 46; Conservative 40; Mismatches 99; Indels 37; Gaps 11;

Qy 24 GVRLLPFCRGKMEIANFPVRPSDVWIVTY-PKSGTSLLOEVVYVLSQGADPDEIGLMI 82
Db 1988 GFNMIPYLRKVLKAISSI-----HSLWQVPYFSKAWQRVLIQPLLSASCPLSPDQ--LSDC 2041

Qy 83 DEQL-PVLEYPOPG-LDIKELTSPRLIKSHLP-----YRPLP-SDLHNGDSKVIYMAR 133
Db 2042 SESLAVLECPVSGDLDLIG--VARQYIQLELPAPAFALACLMLMPHSEKRHOQIKNFLGSC 2099

Qy 134 NPKDLVSVYQFHRSLRTMSYRGTFQFCRRFMND-----KLGYGSWFHEHVQEFWEHRMD 188
Db 2100 DPQ---VLLKQLEEHMTGQLAGFSHQIRSLILNIIINKKEGIIUAKTKIFQMLKQHWAN 2156

Qy 189 SNVLFLKYEDMHRDLVMTVEQLARELVGSCDKAQLEALTEHC 230
Db 2157 TN-----NITELVNYLANDLSLDEASVLITEYSKHC 2187

RESULT 84
US-10-511-937-2546
; Sequence 2546, Application US/10511937
; Publication No. US20060088836A1
; GENERAL INFORMATION:
; APPLICANT: EXPRESSION DIAGNOSTICS, INC.
; APPLICANT: Wohlgenuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; APPLICANT: Prentice, James
; APPLICANT: Morris, Macdonald
```

```
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND MONITORING TRANSPLANT REJECTION
; FILE REFERENCE: 506612000104
; CURRENT APPLICATION NUMBER: US/10/511,937
; CURRENT FILING DATE: 2004-10-19
; PRIOR APPLICATION NUMBER: PCT/US2003/012946
; PRIOR FILING DATE: 2003-04-24
; PRIOR APPLICATION NUMBER: US 10/131,831
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US 10/325,899
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 3117
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2546
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-511-937-2546

Query Match      3.6%; Score 55; DB 6; Length 274;
Best Local Similarity 22.5%; Pred. No. 82;
Matches 40; Conservative 19; Mismatches 59; Indels 60; Gaps 9;

Qy 1 MAESEAEPTSTPGREESKYFEFHGVRLE-----PFCRGKMEEIANFPVRPSDVWIVTYP 54
Db 105 LAQLDGEDPCT--EFRQQLTVHRVHVTFLPHEPP-----PFRPHDVTIVA-- 148

Qy 55 KSGTSLLOEVVYVLSQGADPDEIGLMI 94
Db 149 QLSWDRLOMLCALCHWPGPMSLALYLTDAEAQQLHFVEASPVLAARQDVAVHYVYREG 208

Qy 95 GLDIKELTSPRLIKSHLPYRPLPSDLHNGDSKVIYMARNPDKLVVSY--YQFHRSLR 150
Db 209 PLYPNVQLRNVALAQALTPYVFL-SDI-----DPLPAYSLYDYLREAR 250

RESULT 85
US-11-283-522-35
; Sequence 35, Application US/11283522
; Publication No. US20060088914A1
; GENERAL INFORMATION:
; APPLICANT: Adams, Sean
; APPLICANT: Pan, James
; TITLE OF INVENTION: UCP5
; FILE REFERENCE: P1663R2
; CURRENT APPLICATION NUMBER: US/11/283,522
; CURRENT FILING DATE: 2005-11-18
; PRIOR APPLICATION NUMBER: US/10/270,861
; PRIOR FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US/09/433,622
; PRIOR FILING DATE: 1999-11-02
; PRIOR APPLICATION NUMBER: US 60/110,286
; PRIOR FILING DATE: 1998-11-30
; PRIOR APPLICATION NUMBER: US 60/129,583
; PRIOR FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: US 60/143,886
; PRIOR FILING DATE: 1999-07-15
; NUMBER OF SEQ ID NOS: 36
; SEQ ID NO 35
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Homo Sapien
US-11-283-522-35

Query Match      3.6%; Score 55; DB 7; Length 300;
Best Local Similarity 25.2%; Pred. No. 94;
Matches 30; Conservative 17; Mismatches 44; Indels 28; Gaps 6;

Qy 97 DIIKE-LTSPRLIKSHLPYRPLPSDLHNGDSKVIYMARNPDKLVVSY-----QFHRSL 149
Db 189 DILKEKLDYHLLTDNFPCHFVSA---FGAGFCATVVASPDVVVKTRIMNSPPGOYFSPL 245
```

Qy 150 RTMSYRGTFQE---FCRRPMNDKLGYSWFVHQEFWEHRMDSNVLFKYVEDMHRDLV 204
| : || : | : ||| : || : || : || :
Db 246 DCM:IKMVAEGGPTAFYKGFPSFLRGSW-----NVVMFTYEQLRKALM 290

```

RESULT 86
US-10-511-937-2511
; Sequence 2511, Application US/10511937
; Publication No. US20060088836A1
; GENERAL INFORMATION:
; APPLICANT: EXPRESSION DIAGNOSTICS, INC.
; APPLICANT: Wohlgemuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; APPLICANT: Prentice, James
; APPLICANT: Morris, MacDonald
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND MONITORING TRANSPLANT REJECTION
; FILE REFERENCE: 506612000104
; CURRENT APPLICATION NUMBER: US/10/511.937
; CURRENT FILING DATE: 2004-10-19
; PRIOR APPLICATION NUMBER: PCT/US2003/012946
; PRIOR FILING DATE: 2003-04-24
; PRIOR APPLICATION NUMBER: US 10/131,831
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US 10/325,899
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 3117
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2511
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-511-937-2511

```

	Query Match	3.6%	Score 55;	DB 6;	Length 343;
	Best Local Similarity	27.2%	Pred. No. 1.2e+02;		
	Matches	Conservative	Matches	Indels	Gaps
	25;		28;		
Qy	170 LGYGSWFE-----HVGQFWHRM-DSNVLFKY-----EDMHRDLVTWVEQL	210			
Dd	19 LNEGAWTVSGALMISVPALLGYLGVCRARLPDSELMRKYHSVRQEDLQRVLSRPEAV	78			
Qy	211 ARFLGVSCDKAQLKALTEHCHEQLVDQCCEAEF	242			
Dd	79 AE---VKSFLLTLEAF-----LSRLCCTCEFA	101			

```

RESULT 87
US-11-305-413-13
, Sequence 13, Application US/11305413
, Publication No. US2006009699A1
, GENERAL INFORMATION:
, APPLICANT: Clausen, Henrik
, APPLICANT: Amado, Margerita
, TITLE OF INVENTION: UDP-GALACTOSE: BETA-N-ACETYL-GLUCOSAMINE BETA 1,3
, TITLE OF INVENTION: GALACTOSYLTRANSFERASES, BETAGAL-T5
, FILE REFERENCE: 7188-157
, CURRENT APPLICATION NUMBER: US/11/305,413
, CURRENT FILING DATE: 2005-12-16
, PRIOR APPLICATION NUMBER: US/10/777,828
, PRIOR FILING DATE: 2004-02-12
, PRIOR APPLICATION NUMBER: US/09/831,630
, PRIOR FILING DATE: 2001-05-10
, NUMBER OF SEQ ID NOS: 17
, SOFTWARE: Patentin Ver. 2.0
, SEQ ID NO 13
, LENGTH: 378
, TYPE: PRT
, ORGANISM: Homo sapiens
US-11-305-413-13

```

[illegible]

```

RESULT 88
US-10-196-749-152
; Sequence 152, Application US/10196749
; Publication No. US20060094864A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME
; FILE REFERENCE: P34301C340
; CURRENT APPLICATION NUMBER: US/10/196,749
; CURRENT FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 152
; LENGTH: 541
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-196-749-152

```


Db 241 EKESASAS-----ELHKLQFLLRRVKAEVAT--ELPKKTEVVY--HGMSALQ 286
QY 63 EYVYLVSOGADPD-----EIGLMNIDEQL-----EVL---EYPOPLDIIKELTSP 105
Db 287 KKYKAILMKDLDAPEKETAKVKVLQNILSQLRCKVDHPYLPDGVPEPEP-FEVDGHLTEA 345
QY 106 RLIIKSHLPYRFLPSDLHNGDSKVIYARNPK--DLVVSYVQFHRSLRTWTSYRG 156
Db 346 S-GKLHLLDKLL-AFLYSGHRVLLFSQMTOMLIDLDY-----MDYRG 387

RESULT 95
US-10-505-928-359
; Sequence 359, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; PRIOR FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 359
; LENGTH: 1015
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-359

Query Match 3.6%; Score 54.5; DB 6; Length 1015;
Best Local Similarity 19.5%; Pred. No. 6.8e+02;
Matches 43; Conservative 33; Mismatches 70; Indels 75; Gaps 10;
QY 3 ESEATPTPGFESKYFEFFHGVLPFCRGKMEIANFPVRPSDV-----WIVT-----Y 53
Db 472 ELQNMFGSKEEQS-----LPAGAQAELSDGLQLEVQFSSEEARGYIVTRDPLR 522
QY 54 PKSGTSLQEVVYLV-----SQGADPEIG-----LMNIDEQLPVLEYPQPGLDII 99
Db 523 PEGRRLEDVARLLQVSSAFADVEVLGPATVFKVSANQVNTTE-DVEKATVNDKXL 581
QY 100 KELTSPRLIK----SHLPYRFLPSDLHNGDSK-----VIYMARN 134
Db 582 BETSGLKILQTVGSKSKLFLPPQAEQEDSTKFTALTIVSLACILVLLASGLIVCLR 641
QY 135 PKDLVSYVQFHRSLRTWTSYRG-----TFQEFRRFM 166
Db 642 SSQ-----HRLKELSLGGLGDFGADATAYQELCRQM 674

RESULT 96
US-11-311-778-24
; Sequence 24, Application US/11311778
; Publication No. US20060095988A1
; GENERAL INFORMATION:
; APPLICANT: Fischhoff, David A.
; APPLICANT: Perlak, Frederick J.
; TITLE OF INVENTION: SYNTHETIC PLANT GENES AND METHOD FOR PREPARATION
; FILE REFERENCE: 38-21 (13553) C
; CURRENT APPLICATION NUMBER: US/11/311,778
; CURRENT FILING DATE: 2005-12-19
; PRIOR APPLICATION NUMBER: 08/434,105
; PRIOR FILING DATE: 1995-05-03
; PRIOR APPLICATION NUMBER: 07/959,506
; PRIOR FILING DATE: 1992-10-09
; PRIOR APPLICATION NUMBER: 07/476,661
; PRIOR FILING DATE: 1990-02-12
; PRIOR APPLICATION NUMBER: 07/315,355
; PRIOR FILING DATE: 1989-02-24
; NUMBER OF SEQ ID NOS: 43

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 24
; LENGTH: 1177
; TYPE: PRT
; ORGANISM: artificial
; FEATURE:
; OTHER INFORMATION: synthetic structural gene encoding a fusion protein derived from
; OTHER INFORMATION: B.t.k. HD-1 and B.t.k. HD-73
US-11-311-778-24

Query Match 3.6%; Score 54.5; DB 7; Length 1177;
Best Local Similarity 28.8%; Pred. No. 8.5e+02;
Matches 23; Conservative 8; Mismatches 32; Indels 17; Gaps 4;
QY 7 ETPSTPGFESKYFE---FHGVRLPP---FCRGKME-----EIANFPVRPSDVI 50
Db 1098 EAPSPADYASVYEKSYTDGRRENPCFENRGYRDYTLPLVGVTYKLEYFP-ETDKWI 1156
QY 51 VTYPKSGTSLQEVVYLVSQ 70
Db 1157 EIGETEGTFIVDSVELLME 1176

RESULT 97
US-11-311-778-10
; Sequence 10, Application US/11311778
; Publication No. US20060095988A1
; GENERAL INFORMATION:
; APPLICANT: Fischhoff, David A.
; APPLICANT: Perlak, Frederick J.
; TITLE OF INVENTION: SYNTHETIC PLANT GENES AND METHOD FOR PREPARATION
; FILE REFERENCE: 38-21 (13553) C
; CURRENT APPLICATION NUMBER: US/11/311,778
; CURRENT FILING DATE: 2005-12-19
; PRIOR APPLICATION NUMBER: 08/434,105
; PRIOR FILING DATE: 1995-05-03
; PRIOR APPLICATION NUMBER: 07/959,506
; PRIOR FILING DATE: 1992-10-09
; PRIOR APPLICATION NUMBER: 07/476,661
; PRIOR FILING DATE: 1990-02-12
; PRIOR APPLICATION NUMBER: 07/315,355
; PRIOR FILING DATE: 1989-02-24
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 1178
; TYPE: PRT
; ORGANISM: artificial
; FEATURE:
; OTHER INFORMATION: synthetic structural gene encoding an insecticidal protein derive
; OTHER INFORMATION: d from B.t.k. HD-73
US-11-311-778-10

Query Match 3.6%; Score 54.5; DB 7; Length 1178;
Best Local Similarity 28.8%; Pred. No. 8.5e+02;
Matches 23; Conservative 8; Mismatches 32; Indels 17; Gaps 4;
QY 7 ETPSTPGFESKYFE---FHGVRLPP---FCRGKME-----EIANFPVRPSDVI 50
Db 1099 EAPSPADYASVYEKSYTDGRRENPCFENRGYRDYTLPLVGVTYKLEYFP-ETDKWI 1157
QY 51 VTYPKSGTSLQEVVYLVSQ 70
Db 1158 EIGETEGTFIVDSVELLME 1177

RESULT 98
US-11-311-778-12
; Sequence 12, Application US/11311778
; Publication No. US20060095988A1
; GENERAL INFORMATION:
; APPLICANT: Fischhoff, David A.
; APPLICANT: Perlak, Frederick J.

GenCore version 5.1.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 17, 2006, 10:59:35 ; Search time 40 Seconds
(without alignments)
683.139 Million cell updates/sec

Title: US-10-768-158-2

Perfect score: 1530

Sequence: 1 MAESEAEPTSPGPESEKYF.....FDLVYKQKMGKDLTFDFYL 284

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

PIR 80:*

1: Pirl:*

2: Pirl:*

3: Pirl:*

4: Pirl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length	ID	Description
1	495	32.4	298	2 S28183	aryl sulfotransfer
2	472	30.8	295	2 A55451	aryl sulfotransfer
3	467	30.5	296	2 JC7282	hydroxyarylamine s
4	458	29.9	291	2 S10329	aryl sulfotransfer
5	456	29.8	295	2 S52791	aryl sulfotransfer
6	456	29.8	295	2 JC5249	aryl sulfotransfer
7	456	29.8	295	2 G01843	aryl sulfotransfer
8	454	29.7	295	2 S52794	phenol-sulfating p
9	451	29.5	295	2 S52794	aryl sulfotransfer
10	451	29.5	295	2 JC5248	aryl sulfotransfer
11	449	29.3	295	2 JC5253	aryl sulfotransfer
12	447	29.2	296	2 JC7283	hydroxyarylamine s
13	446.5	29.2	299	2 JC5884	thyroid hormone su
14	446	29.2	295	2 G02924	phenol sulfotransf
15	446	29.2	295	2 S29045	estrone sulfotrans
16	444	29.0	295	2 S52399	aryl sulfotransfer
17	440.5	28.8	296	2 JC5885	thyroid hormone su
18	435	28.4	295	2 I73679	estrogen sulfotran
19	434	28.4	294	2 JC2229	estrogen sulfotran
20	432.5	28.3	299	2 JE0186	sulfotransferase (
21	430	28.1	295	2 I53296	testis-specific es
22	430	28.1	295	2 I56606	estrogen sulfotran
23	429	28.0	295	2 A41930	estrone sulfotrans
24	419.5	27.4	304	2 A49098	N-hydroxyarylamine
25	413	27.0	295	2 JE0197	phenol sulfotransf
26	412	26.9	294	2 JC5000	aryl sulfotransfer
27	408.5	26.7	296	2 A44011	adrenocortical est
28	393.5	25.7	301	2 JM0078	amine sulfotransf
29	385.5	25.2	287	2 JC4531	alcohol sulfotrans

30	380	24.8	338	2	JE0196	hydroxysteroid sul
31	376.5	24.6	287	2	A54026	alcohol sulfotrans
32	371	24.2	286	2	JE0152	alcohol sulfotrans
33	361.5	23.6	285	2	I65760	alcohol sulfotrans
34	359.5	23.5	284	2	I60190	alcohol sulfotrans
35	356.5	23.3	285	2	T10086	alcohol sulfotrans
36	355.5	23.2	285	1	I13848	alcohol sulfotrans
37	353.5	23.1	284	2	I52849	alcohol sulfotrans
38	351.5	23.0	284	2	A33569	alcohol sulfotrans
39	338	22.1	282	2	A26136	senescence marker
40	303	19.8	311	2	B40216	flavonol 3'-sulfot
41	302	19.7	329	2	T47448	sulfotransferase-1
42	297.5	19.4	324	2	B84452	probable steroid s
43	284.5	18.6	323	2	T47447	sulfotransferase-1
44	280	18.3	325	2	T07833	probable steroid s
45	278.5	18.2	346	2	E86319	probable flavonol
46	277.5	18.1	326	2	A84452	probable steroid s
47	276.5	18.1	323	2	T07831	probable steroid s
48	275.5	18.0	320	2	A40216	flavonol 4'-sulfot
49	267	17.5	350	2	H96768	protein flavonol s
50	265.5	17.4	324	2	T07832	probable steroid s
51	262.5	17.2	326	2	F86407	probable sulfotran
52	259.5	17.0	302	2	S69188	probable flavonol
53	254.5	16.6	296	2	JC7921	cytosolic sulfotra
54	249.5	16.3	338	2	A96769	protein flavonol s
55	240	15.7	333	2	A84523	probable steroid s
56	235.5	15.4	314	2	T06012	hypothetical prote
57	224.5	14.7	331	2	H84451	probable steroid s
58	161	10.5	280	2	A95936	probable alcohol s
59	110	7.2	249	2	E84072	hypothetical prote
60	109.5	7.2	761	2	G70393	conserved hypotet
61	106.5	7.0	326	2	B70958	hypothetical prote
62	101	6.6	5107	2	T29144	partial CDS - Caen
63	100.5	6.6	1424	2	T03851	thyroid hormone re
64	96	6.3	481	2	T27665	hypothetical prote
65	94.5	6.2	678	2	B64592	hypothetical prote
66	93	6.1	363	2	S18765	Sd protein - fruit
67	92.5	6.0	500	2	T03387	polyamine oxidase
68	92	6.0	759	2	T39468	anthranilate synth
69	91.5	6.0	1434	2	G71232	hypothetical prote
70	91	5.9	285	2	S49879	hypothetical narbo
71	91	5.9	285	2	S49898	hypothetical narbo
72	91	5.9	508	2	A71453	hypothetical prote
73	91	5.9	759	2	T46566	anthranilate synth
74	90.5	5.9	882	2	A57169	(heparan sulfate) -
75	90	5.9	435	2	D72379	hypothetical prote
76	89	5.8	332	2	B99791	hypothetical prote
77	88.5	5.8	882	2	A42855	N-heparan sulfate
78	88	5.8	604	2	C57784	MTG8 protein aplic
79	88	5.8	1752	4	A57784	AMU1/MTG8 mutant f
80	88	5.8	1762	2	T00458	hypothetical prote
81	87.5	5.7	254	2	B84071	hypothetical prote
82	87.5	5.7	308	2	C95935	probable sulfotran
83	87.5	5.7	466	2	H98017	conserved hypotet
84	87.5	5.7	883	2	A49733	(heparan sulfate) -
85	87.5	5.7	976	2	C84500	probable retroelem
86	87.5	5.7	1455	2	E75199	DNA-directed DNA p
87	86.5	5.7	434	2	A95150	conserved hypotet
88	86	5.6	836	1	JDVLD	DNA-directed DNA p
89	86	5.6	1452	2	T29107	polymorphic antige
90	85.5	5.6	352	2	A9298	conserved hypotet
91	85.5	5.6	1462	2	T42639	glucocorticoid rec
92	85	5.6	404	2	A34919	type II site-speci
93	85	5.6	463	2	S75889	hypothetical prote
94	85	5.6	607	1	J01535	hemagglutinin - ph
95	85	5.6	1105	2	S76557	carbamoyl-phosphat
96	84.5	5.5	782	2	D84514	dynamun-like prote
97	84	5.5	2747	2	B49132	fat facets (faf) s
98	83.5	5.5	419	2	C81744	ATP-dependent Clp
99	83.5	5.5	778	2	A86928	probable ATP-depen
100	83.5	5.5	1071	2	T22327	hypothetical prote

QY 243 LP-VGRGRVGLWKDIPTVSMNEKFDLVYKQKMGKCDLTF 280
Db 254 -PFMRKGMAGDWKTTFTVAQNERFDADYAEKMGAGCSLSF 291

RESULT 3
JC7282

hydroxyarylamine sulfotransferase (EC 2.8.2.-) 2 - rat
N:Alternate names: SULT1C2 sulfotransferase
C:Species: Rattus norvegicus (Norway rat)
C:Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 09-Jul-2004
C:Accession: JC7282
R:Li, X.; Joehnk, C.; Hartmann, D.; Schestag, F.; Kroemer, W.; Gieselmann, V.
Biochem. Biophys. Res. Commun. 272, 242-250, 2000
A:Title: Enzymatic properties, tissue-specific expression, and lysosomal location of two
A:Reference number: JC7282
A:Contents: Kidney
A:Accession: JC7282
A:Molecule type: mRNA
A:Residues: 1-236 <L1A>
A:Cross-references: UNIPROT:Q9WU8; UNIPARC:UPI00001360FC; GB:AJ238391
C:Genetics:
A:Gene: sult1c2
C:Superfamily: alcohol sulfotransferase
C:Keywords: kidney; sulfotransferase

Query Match 30.5%; Score 467; DB 2; Length 296;
Best Local Similarity 35.1%; Pred. No. 9.7e-33;
Matches 100; Conservative 58; Mismatches 105; Indels 22; Gaps 5;

QY 16 ESKYFEHGVRLPPFCRGMKEIANFPVRPSDWMIVTPYKSGTSLQEVVYLSQGAPDP 75
Db 10 QTKLEVAGIPLQAPTVDNWSQIQTFKAKPDLLICTYPKSGTTWIEIVDMIEQGDVE 69

QY 76 EIGLMNIDEQLVLEY---POP-GLDIIKELTSPLRIKSHLPYRFLPSDLHNGDSKVIYM 131
Db 70 KCORTIIQHRHFFIEWARPPQSGVDKXANAMPAPRILRTHLPQLLPPSFTWNCKFLV 129

QY 132 ARNPDLVVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHGVQEFWEHRMDSNV 191
Db 130 ARNADCVSYHYFRMSQVLPDPTWNEYFTFNGKVSWSGDFHVKGWWEIRDYQI 189

QY 192 LFLKYEDMHRDLVMTVEQLARFLGVS-----CDKQAQLEALTEHCQLVDQCCNABALP-- 244
Db 190 LFLFYEDVKRDKPEIQKVMQFMGKNLDEVVVDKIVLETSFEKMKES--NPMTNRSTVPKS 247

QY 245 -----VGRGRVGLWKDIPTVSMNEKFDLVYKQKMGKCDLTF 280
Db 248 VLDSQISPFMRKGTVDGWNKHFTVAQNDRFDBIYKQKMGGTSLNF 292

RESULT 4
S10329

aryl sulfotransferase (EC 2.8.2.1) IV - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C:Accession: S10329; A44932
R:Ozawa, S.; Nagata, K.; Gong, D.; Yamazoe, Y.; Kato, R.
Nucleic Acids Res. 18, 4001, 1990
A:Title: Nucleotide sequence of a full-length cDNA (pST-1) for aryl sulfotransferase from
A:Reference number: S10329; MUID:90326537; PMID:2374726
A:Accession: S10329
A:Molecule type: mRNA
A:Residues: 1-291 <OZA>
A:Cross-references: UNIPROT:P17988; UNIPARC:UPI0000000B75; EMBL:X52883; NID:g55760; PIDN:
Riferokun, T.; Etheredge, J.L.; Norton, T.R.; Carter, H.A.; Chung, K.H.; Birckbichler, H.
Cancer Res. 52, 4779-4786, 1992
A:Title: Characterization of a complementary DNA for rat liver aryl sulfotransferase IV
used hepatocarcinogenesis.
A:Reference number: A44932; MUID:92379794; PMID:1511441
A:Accession: A44932
A:Status: preliminary

A:Molecule type: mRNA; protein
A:Residues: 7-291 <YER>
A:Cross-references: UNIPARC:UPI000017088C; GB:X68640; GB:S42994; NID:g55764; PIDN:CAA486
A:Experimental source: liver
A:Note: sequence extracted from NCBI backbone (NCBIN:112171, NCBIP:112178)
C:Superfamily: alcohol sulfotransferase
C:Keywords: sulfotransferase

Query Match 29.9%; Score 458; DB 2; Length 291;
Best Local Similarity 34.9%; Pred. No. 5.7e-32;
Matches 96; Conservative 57; Mismatches 104; Indels 18; Gaps 4;

QY 24 GVRLPFFCRGMKEIANFPVRPSDWMIVTPYKSGTSLQEVVYLSQGAPDPDEIGLMNID 83
Db 13 GIPLIKYFAETIGPLQNTAMPDDLLISTYPKSGTWTMSDEILDIMYQGGKLEKCGRAPYI 72

QY 84 EQLPVLEYPOP---GLDIIKELTSPLRIKSHLPYRFLPSDLHNGDSKVIYMARNP KDIV 139
Db 73 ARVPFLEFKCPGVPSGLETLETPAPRLLKTHLPQLSLPQSLLDQKVKVIYIARNAKDVV 132

QY 140 VSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHGVQEFWEHRMDSNVFLKYEDM 199
Db 133 VSYNYFNMAKLIHPDPGTWDSFLENFMDEGVSYSWYQHKVWELRHTHPVLVLYFYEDI 192

QY 200 HRDLVTWVEQLARFLGVS CDKQAQLEALTEHC--HQLVDQC--CNAEALP-----V 245
Db 193 KENPKREIKKILEFLGRSLPEETVDSIVVHTSFKKMKENCMTNYTTIPTIEMDHNVSPPM 252

QY 246 GRGRVGLWKDIPTVSMNEKFDLVYKQKMGKCDLTF 280
Db 253 RKGTTGDWKTFTVAQNERFDAHYAKTWTDCDFK 287

RESULT 5
S52791

aryl sulfotransferase (EC 2.8.2.1) - human
C:Species: Homo sapiens (man)
C:Date: 19-May-1995 #sequence_revision 21-Jul-1995 #text_change 18-Jun-1999
C:Accession: S52791
R:Ozawa, S.; Nagata, K.; Shimada, M.; Ueda, M.; Tsuzuki, T.; Yamazoe, Y.; Kato, R.
submitted to the EMBL Data Library, March 1994
A:Description: Primary structures and properties of two related aryl sulfotransferases
A:Reference number: S52791
A:Accession: S52791
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-295 <OZA>
A:Cross-references: UNIPARC:UPI000016806E; EMBL:X78282; NID:g758594; PIDN:CAA55088.1; PI:
C:Superfamily: alcohol sulfotransferase
C:Keywords: sulfotransferase

Query Match 29.8%; Score 456; DB 2; Length 295;
Best Local Similarity 33.7%; Pred. No. 8.6e-32;
Matches 94; Conservative 58; Mismatches 101; Indels 26; Gaps 4;

QY 24 GVRLPFFCRGMKEIANFPVRPSDWMIVTPYKSGTSLQEVVYLSQGAPDPDEIGLMNID 83
Db 17 GVLLIKYFAEALGPLQSFQARPDLLISTYPKSGTTWVSQILDIMYQGGDLEKCHRAPTF 76

QY 84 EQLPVLEYPOP---GLDIIKELTSPLRIKSHLPYRFLPSDLHNGDSKVIYMARNP KDIV 139
Db 77 MRVPFLEFKCPGVPSGMETLKNTPAPRLLKTHLPQLLDPQLLDQKVKVIYIARNAKDOVA 136

QY 140 VSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHGVQEFWEHRMDSNVFLKYEDM 199
Db 137 VSYTHFYHMAKVYPHPGTWESFLEKFMAGEVSYSWYQHKVWELRHTHPVLVLYFYEDM 196

QY 200 HRDLVTWVEQLARFLGVS CDKQAQLEALTEHC-----HQLVDQCCNAEA 242
Db 197 KENPKREIQKILEFVGRSLPEETVDSIVVHTSFKKMKKNPMTNYTTVRRFRFMDHSIS--- 253

QY 243 LP-VGRGRVGLWKDIPTVSMNEKFDLVYKQKMGKCDLTF 280
Db 248 VLDSQISPFMRKGTVDGWNKHFTVAQNDRFDBIYKQKMGGTSLNF 292

Db 254 -PFMRKGMAGDWKTTFTTVAQNERFDADYAEKMGCSLSF 291

RESULT 6
JC5249
aryl sulfotransferase (EC 2.8.2.1) 2 - human
N:Alternate names: phenol sulfotransferase; sulfokinase
C:Species: Homo sapiens (man)
C:Date: 13-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: JC5249
R:Dooley, T.P.; Huang, Z.
Biochem. Biophys. Res. Commun. 228, 134-140, 1996
A:Title: Genomic organization and DNA sequences of two human phenol sulfotransferase genes
A:Reference number: JC5248; MUID:97069665; PMID:8912648
A:Accession: JC5249
A:Molecule type: mRNA
A:Residues: 1-295 <DOO>
A:Cross-references: UNIPROT:P50226; UNIPARC:UPI0000136229; GB:U76619; NID:g1666920; PIDN:
C:Comment: This enzyme catalyzes the transfer of the sulfonate group from phosphoadenosine
C:Genetics:
A:Gene: stp2
A:Map position: 16p12.1-pl1.2
C:Superfamily: alcohol sulfotransferase
C:Keywords: sulfotransferase

Query Match 29.8%; Score 456; DB 2; Length 295;
Best Local Similarity 33.7%; Pred. No. 8.6e-32;
Matches 94; Conservative 58; Mismatches 101; Indels 26; Gaps 4;

Qy 24 GVRLLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSLQEVVYLSQAGDPDEIGLMNID 83
Db 17 GVPLIKYFAEALGPLQSFQARPDLLISTYPKSGTTWVSQILDMIYQGGDLEKCHRAPIF 76

Qy 84 EQLPVLEYPOP-GLDIIKELTSPLIKSHLPYRFLPSDLHNGDSKVIYMARNPXDLV 139
Db 77 MRVPFLFKVPGIPSGMETLKNTPAPRLKTHLPLALLPQTLQDKQKVVYVARNAKDVA 136

Qy 140 VSYQFHRSRLTMSYRGTFQFCRRFMNDKLGYSWFEHVEFWEHRMDSNVLFKLYEDM 199
Db 137 VSYHYFHYMAKVYPHPGTWESFLEKFMAGEVSGYSWYQHVQEWELSRTHPVLVLYFYEDM 196

Qy 200 HRDLVTWVEQLARFLGVSCDKAQLEALTEHC-----HQLVDQCCNAEA 242
Db 197 KENPKREIQKILEFVGRSLPBEETVDLMEHTSFPEKMKKNPMTNVTYVRRPFMDHSIS--- 253

Qy 243 LP-VGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTF 280
Db 254 -PFMRKGMAGDWKTTFTTVAQNERFDADYAEKMGCSLSF 291

RESULT 7
G01843
aryl sulfotransferase - human
C:Species: Homo sapiens (man)
C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 09-Jul-2004
C:Accession: G01843
R:Zhu, X.
submitted to the EMBL Data Library, May 1995
A:Reference number: G08606
A:Accession: G01843
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-295 <ZHU>
A:Cross-references: UNIPROT:P50226; UNIPARC:UPI000015486D; EMBL:U28169; NID:g881502; PID
C:Superfamily: alcohol sulfotransferase

Query Match 29.8%; Score 456; DB 2; Length 295;
Best Local Similarity 33.7%; Pred. No. 8.6e-32;
Matches 94; Conservative 58; Mismatches 101; Indels 26; Gaps 4;

Qy 24 GVRLLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSLQEVVYLSQAGDPDEIGLMNID 83
Db 17 GVPLIKYFAEALGPLQSFQARPDLLISTYPKSGTTWVSQILDMIYQGGDLEKCHRAPIF 76

Qy 84 EQLPVLEYPOP-GLDIIKELTSPLIKSHLPYRFLPSDLHNGDSKVIYMARNPXDLV 139
Db 77 MRVPFLFKVPGIPSGMETLKNTPAPRLKTHLPLALLPQTLQDKQKVVYVARNAKDVA 136

Qy 140 VSYQFHRSRLTMSYRGTFQFCRRFMNDKLGYSWFEHVEFWEHRMDSNVLFKLYEDM 199
Db 137 VSYHYFHYMAKVYPHPGTWESFLEKFMAGEVSGYSWYQHVQEWELSRTHPVLVLYFYEDM 196

Qy 200 HRDLVTWVEQLARFLGVSCDKAQLEALTEHC-----HQLVDQCCNAEA 242
Db 197 KENPKREIQKILEFVGRSLPBEETVDLMEHTSFPEKMKKNPMTNVTYVRRPFMDHSIS--- 253

Qy 243 LP-VGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTF 280
Db 254 -PFMRKGMAGDWKTTFTTVAQNERFDADYAEKMGCSLSF 291

RESULT 8

157945

phenol-sulfating phenol sulfotransferase - human

C:Species: Homo sapiens (man)

C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004

C:Accession: I57945

R:Wilborn, T.W.; Comer, K.A.; Dooley, T.P.; Reardon, I.M.; Heinrikson, R.L.; Falany, C.N.

Mol. Pharmacol. 43, 70-77, 1993

A:Title: Sequence analysis and expression of the cDNA for the phenol-sulfating form of h

A:Reference number: I57945; MUID:93140712; PMID:8423770

A:Accession: I57945

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-295 <RES>

A:Cross-references: UNIPROT:P50225; UNIPARC:UPI000016AEF7; GB:L19999; NID:g307342; PIDN:

C:Superfamily: alcohol sulfotransferase

Query Match

29.7%; Score 454; DB 2; Length 295;

Best Local Similarity 34.2%; Pred. No. 1.3e-31;

Matches 94; Conservative 56; Mismatches 107; Indels 18; Gaps 3;

Qy 24

GVRLLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSLQEVVYLSQAGDPDEIGLMNID 83

Db 17 GVPLIKYFAEALGPLQSFQARPDLLISTYPKSGTTWVSQILDMIYQGGDLEKCHRAPIF 76

Qy 84

EQLPVLEYPOP-GLDIIKELTSPLIKSHLPYRFLPSDLHNGDSKVIYMARNPXDLV 139

Db 77 MRVPFLFKVPGIPSGMETLKNTPAPRLKTHLPLALLPQTLQDKQKVVYVARNAKDVA 136

Qy 140

VSYQFHRSRLTMSYRGTFQFCRRFMNDKLGYSWFEHVEFWEHRMDSNVLFKLYEDM 199

Db 137 VSYHYFHYMAKVYPHPGTWDSFLEKFMAGEVSGYSWYQHVQEWELSRTHPVLVLYFYEDM 196

Qy 200

HRDLVTWVEQLARFLGVSCDKAQLEALTEHC-----HQLVDQCCNAEA 245

Db 197 KENPKREIQKILEFVGRSLPBEETVDFMVQHTSFPEKMKKNPMTNVTYVQEPFMDHSISPFM 256

Qy 246

GRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTF 280

Db 257 RKGMDAGDWKTTFTTVAQNERFDADYAEKMGCSLSF 291

RESULT 9

S52794

aryl sulfotransferase (EC 2.8.2.1) - human

C:Species: Homo sapiens (man)

C:Date: 19-May-1995 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004

C:Accession: S52794

R:Ozawa, S.; Nagata, K.; Shimada, M.; Ueda, M.; Tsuzuki, T.; Yamazoe, Y.; Kato, R.

submitted to the EMBL Data Library, March 1994

A:Description: Primary structures and properties of two related aryl sulfotransferases

A:Reference number: S52791

A:Accession: S52794

A:Status: preliminary

A:Molecule type: mRNA

QY 246 GRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 257 RKGAGDWKTTFTVAQNERFDADYAEKMGCSLSF 291

RESULT 12
JC7283
hydroxyarylamine sulfotransferase (EC 2.8.2.2-) 2A - rat
N:Alternate names: SULT1C2A sulfotransferase
C:Species: Rattus norvegicus (Norway rat)
C:Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 02-Sep-2000
C:Accession: JC7283
R:Li, X.; Joehnk, C.; Hartmann, D.; Schestag, F.; Kroemer, W.; Gieselmann, V.
Biochem. Biophys. Res. Commun. 272, 242-250, 2000
A:Title: Enzymatic properties, tissue-specific expression, and lysosomal location of two
A:Reference number: JC7282
A:Contents: Kidney
A:Accession: JC7283
A:Molecule type: mRNA
A:Residues: 1-296 <L1A>
C:Cross-references: UNIPARC:UPI0000175833; GB:AJ238392
C:Genetics:
A:Gene: sult1c2a
A:Superfamily: alcohol sulfotransferase
C:Keywords: kidney; sulfotransferase

Query Match 29.2%; Score 447; DB 2; Length 296;
Best Local Similarity 34.4%; Pred. No. 5.2e-31;
Matches 98; Conservative 57; Mismatches 108; Indels 22; Gaps 5;

QY 16 ESKYFEFGVRLPPRCRGMIEANFPVRPSDWMIVTPKSGTSLLOEVVYVLSQADPD 75
Db 10 QTKLEVAGIPLRSTVDNWSQIQTFKAPDPLLCTYKSGTWTWQIENVNMIQONGDVE 69
QY 76 EIGLMNIDEQPLVLEY--POP-GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVYIM 131
Db 70 KCQRTIIQHRPFFIWARPPQSGVDKANAMPAPILTHLPILQILLPSPFWTNCKYLYV 129
QY 132 ARNPKDLVSVYQFHRSLRTMSYRGTQFQFCRRFMNDKLGYSWFHVEHMRDMSNV 191
Db 130 ARNAKDCMVSFYHFYRMCQVLPNPGTWNEYPETFTFNGKVSVMGSCPDHVKGWWEIRDYQI 189
QY 192 LFLKYEDHRLDVTMVEQLAEFLVGS-----CDKAQLEALTEHCHQLVDQCCNAEALP-- 244
Db 190 LFLFYEDMKRDPKREIQKVMQFMGNLDEEVVDVKLVLETSEKMD--NPLTNFTSTIPKT 247
QY 245 -----VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 248 IMDQISFPFRKGI VGDWKQHFTVAQNERFDEIYEQKMDGTSLSNF 292

RESULT 13
JC5884
thyroid hormone sulfotransferase (EC 2.8.2.2-) B1 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 11-Mar-1998 #sequence_revision 11-Mar-1998 #text_change 09-Jul-2004
C:Accession: JC5884
R:Fujita, K.; Nagata, K.; Ozawa, S.; Sasano, H.; Yamazoe, Y.
J. Biochem. 122, 1052-1061, 1997
A:Title: Molecular cloning and characterization of rat ST1B1 and human ST1B2 cDNAs, enco
A:Reference number: JC5884; MUID:98104061; PMID:9443824
A:Accession: JC5884
A:Molecule type: mRNA
A:Residues: 1-299 <FUJ>
C:Cross-references: UNIPROT:P52847; UNIPARC:UPI00001361C1; DDBJ:D89375; NID:g2826147; PI
A:Experimental source: liver
C:Comment: This enzyme mediates the transfer of SO3- from 3'-phosphoadenosine 5'-phospho
C:Superfamily: alcohol sulfotransferase
C:Keywords: sulfotransferase

Query Match 29.2%; Score 446.5; DB 2; Length 299;
Best Local Similarity 36.0%; Pred. No. 5.8e-31;

Matches 95; Conservative 52; Mismatches 98; Indels 19; Gaps 3;
QY 36 BEIANFPVRSDWMIVTPKSGTSLLOEVVYVLSQADPDIEIGLMNIDEQPLVLEYPOP- 94
Db 29 EKIEFQSRPCDIVIPTYPKSGTWTWSEIVDMVLNDGNVEKCKRDVITSKYPMLEQNVPG 88
QY 95 ---GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVYIMARNPKDLVSVYQFHRSLR 150
Db 89 ARRSGBELLKTPSPRIIKTHLPIDLLPKSFDWNKCKMIVLARNGKDVAVSYIYFDLMNN 148
QY 151 TMSYRGTQFQFCRRFMNDKLGYSWFHVEHMRDMSNVFLKYEDHMRDLVTMVEQL 210
Db 149 IQPLPGTWEEYLEKFLAGNAVYGSWFDHVKSWKEKRGHPILFLYEDLKKNPKEIKKI 208
QY 211 ARFLGVSCDRAQLAEALTEHCHQLV---DQCCNAEALP-----VGRGRVGLWKDI 256
Db 209 ANFLDKLTDEHTLERIVHHTSFYEMKONPLVNYTHLPTEIMDHSKSPFMRKGVVGDWKNY 268
QY 257 FTVSMNEKFDLVYKQKMGKCDLTF 280
Db 269 FTMQSEKFDIYKKLSGTTLEF 292

RESULT 14
G02924
phenol sulfotransferase subunit - crab-eating macaque (fragment)
C:Species: Macaca fascicularis (crab-eating macaque)
C:Date: 21-Dec-1996 #sequence_revision 08-Jun-1997 #text_change 09-Jul-2004
C:Accession: G02924
R:Ogura, K.
submitted to GenBank, May 1996
A:Reference number: H01975
A:Accession: G02924
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-295 <OGU>
C:Cross-references: UNIPROT:P52846; UNIPARC:UPI000013622B; GB:D85514; NID:g1339922; PIDN
C:Superfamily: alcohol sulfotransferase

Query Match 29.2%; Score 446; DB 2; Length 295;
Best Local Similarity 33.8%; Pred. No. 6.3e-31;
Matches 93; Conservative 56; Mismatches 108; Indels 18; Gaps 3;

QY 24 GVRLPPECRGMIEANFPVRPSDWMIVTPKSGTSLLOEVVYVLSQADPDIEIGLMNID 83
Db 17 GVPLIKYFAEALGPLQSFRAFPDPLLSTTPKSGTWTWVSQLDMWYQGDLEKCRAPIF 76
QY 84 EQLFVLEYPOP---GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVYIMARNPKDLV 139
Db 77 MRVPFLEFKVPGIFSGMETLKDTPAPRLKTHLPALLPQTLLOQKVKVYVARNKQVA 136
QY 140 VSVYQFHRSLRTMSYRGTQFQFCRRFMNDKLGYSWFHVEHMRDMSNVFLKYEDM 199
Db 137 VSVYHFYHMAKVHPEPGTWDSFLEKFMAGEVSYSGYQHVOEWMLSHTHPVLFLYEDM 196
QY 200 HRDLVTMVEQLAEFLVGSCKAQLAEALTEHC---HQLVDQCCNAEALP-----V 245
Db 197 KENPKREIWKILEFVGRSLPEETVDLMVQHTSFEMKKNPNMANYTTTTPQELMDHSISPFM 256
QY 246 GRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 257 RKGMTGDKWKTFTVAQNEHFDVDYAEKMGCSLSF 291

RESULT 15
S29045
estrone sulfotransferase (EC 2.8.2.4) - bovine
N:Alternate names: estrone sulfotransferase
C:Species: Bos primigenius taurus (cattle)
C:Date: 25-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C:Accession: S29045; S18222
R:Naah, A.R.; Glenn, W.K.; Moore, S.S.; Kerr, J.; Thompson, A.R.; Thompson, E.O.P.
Aust. J. Biol. Sci. 41, 507-516, 1988

Matches 95; Conservative 56; Mismatches 116; Indels 18; Gaps 3;


```
QY 81 NIDEOLPVLEY--POP---GLDIIKELTSPRLIKSHLPVRFPLPSDLHNGDSKVIYMARNP 135
Db 82 NYDRHPFIEWTLPSPLNSGLDLANKMPSPTLKTHLPVHMLPPSPFWKENSKIILYVARNA 141
QY 136 KDLVVSYYQPHRSRLTMSYRGTFQBFRCRRFMNDKLGYSWFEHVOEFWEHRMDSNVLFJK 195
Db 142 KDLVVSYYVFSRMNKLDPDGTGLGSIYEQKAGKVLGWSYDHVKGWMDVKDQHRILYLF 201
QY 196 YEDMRDLVTWVEQLARFLGVSQCDKAQLEALTEHCHQLV---DQCCNAELP----- 244
Db 202 YEDMKEDPKRETKIAKLEKDISEVLNKIYHTSFDMVKENPMANYTTLFPSSIMDSHI 261
QY 245 ---VGRGVGLWKDQJFTVSMNEKFDLVYKQMGKCDLTF 280
Db 262 SPFMKMGPDGWNKYFTVAQSDDFEDIRRMAGSNIYF 300

RESULT 25
JEO197
phenol sulfotransferase (EC 2.8.2.1) - mouse
C:Species: Mus musculus (house mouse)
C>Date: 21-Aug-1998 #sequence_revision 21-Aug-1998 #text_change 09-Jul-2004
C:Accession: JEO197; JG0188
R:Sakakibara, Y.; Yanagisawa, K.; Takami, Y.; Nakayama, T.; Suiko, M.; Liu, M.C.
Biochem. Biophys. Res. Commun. 247, 681-686, 1998
A>Title: Molecular cloning, expression, and functional characterization of novel mouse
A:Reference number: JEO196; MUID:98321187; PMID:9647753
A:Accession: JEO197
A:Molecule type: mRNA
A:Residues: 1-295 <SAK>
A:Cross-references: UNIPROT:Q35401; UNIPARC:UPI00000566C3; GB:AF026073; NID:g2570899; PI
R:Liu, M.; Sakakibara, Y.; Liu, C.
Biochem. Biophys. Res. Commun. 254, 65-69, 1999
A>Title: Bacterial expression, purification, and characterization of a novel mouse sulfo
A:Reference number: JG0188; MUID:99121190; PMID:9920733
A:Accession: JG0188
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-295 <LIU>
A:Cross-references: UNIPARC:UPI00000566C3; GB:AF026073
A:Comment: This enzyme catalyzes the transfer of a sulfonate group from the active sulfa
P.
C:Superfamily: alcohol sulfotransferase
C:Keywords: transferase

Query Match 27.0%; Score 413; DB 2; Length 295;
Best Local Similarity 31.7%; Pred. No. 4.5e-28;
Matches 83; Conservative 63; Mismatches 98; Indels 18; Gaps 2;

QY 37 ETANFPVRPSDVWITYPKSGTSLQEVVLYVLSQGDADPEIGLNMNIDQLPVLEYQPQL 96
Db 30 QVESPEARDDLLISTYPKSGTWTSEILDLLYNNGDAEKCKRDAIYKRVPMELIIPGI 89
QY 97 ----DIKELTSPRLIKSHLPVRFPLPSDLHNGDSKVIYMARNPDLVVSYYQPHRSRLTM 152
Db 90 TNGVEMLNMPSPRIVKTHLPVQLLPSSFWKMDCKLIIVARNAKDVVVSYYFYQWAKIH 149
QY 153 SYRGTFQBFRCRRFMNDKLGYSWFEHVOEFWEHRMDSNVLFJKYEDMRDLVTWVEQLAR 212
Db 150 PEPGTWEEFLEKFMAGQVSFGPWYDHVKSWEKREYRILYLFYEDMKENPKCEIQILK 209
QY 213 FLGVSCDKAQLAEALTEHCH-----QLVDQCCNAELPVGGRVGLMKDIFT 258
Db 210 YLEKDIPEILNKILYHSFSVMKNPNSANYTTMMKEEMDHSVSPFMKRGISGDWKNQFT 269
QY 259 VSMNEKFDLVYKQMGKCDLTF 280
Db 270 VAQYEKFEEDYVKMEDSTLKP 291

RESULT 26
JG5000
aryl sulfotransferase (EC 2.8.2.1), phenol-preferring - bovine
```

```
C:Species: Bos primigenius taurus (cattle)
C>Date: 13-Dec-1996 #sequence_revision 21-Jan-1997 #text_change 09-Jul-2004
C:Accession: JG5000; S59269; S59070
R:Henry, T.; Kliewer, B.; Palmatier, R.; Ulphani, J.S.; Beckmann, J.D.
Gene 174, 221-224, 1996
A>Title: Isolation and characterization of a bovine gene encoding phenol sulfotransferase
A:Reference number: JG5000; MUID:97045816; PMID:8890738
A:Accession: JG5000
A:Molecule type: DNA
A:Residues: 1-294 <HEN>
A:Cross-references: UNIPROT:P50227; UNIPARC:UPI000013622A; GB:U34753; NID:g1276632; PIDN
R:Schaus, S.J.; Henry, T.; Palmatier, R.; Halvorson, L.; Dannenbring, R.; Beckmann, J.D.
Biochem. J. 311, 209-217, 1995
A>Title: Characterization of bovine tracheobronchial phenol sulphotransferase cDNA and
A:Reference number: S59070; MUID:96003918; PMID:7575456
A:Accession: S59269
A:Molecule type: mRNA
A:Residues: 1-294 <SCH>
A:Cross-references: UNIPARC:UPI000013622A; GB:U35253; NID:g1006836; PIDN:AAA85510.1; PID
A:Accession: S59070
A:Molecule type: protein
A:Residues: 23-36738-47; 79-96; 134-144; 207-226; 234-248 <SCW>
A:Cross-references: UNIPARC:UPI0000175829; UNIPARC:UPI000017582A; UNIPARC:UPI000017582B;
C:Comment: This enzyme belongs to a family of cytosolic enzymes that catalyze the transf
C:Genetics:
A:Gene: pst
A:Introns: 50/1; 93/3; 124/3; 167/1; 198/3; 259/1
C:Superfamily: alcohol sulfotransferase
C:Keywords: sulfotransferase

Query Match 26.9%; Score 412; DB 2; Length 294;
Best Local Similarity 33.1%; Pred. No. 5.4e-28;
Matches 91; Conservative 56; Mismatches 108; Indels 20; Gaps 4;

QY 17 SKYFEFGVRLPPFCRGKMEBIANFPVRPSDVWITYPKSGTSLQEVVLYVLSQGDADPE 76
Db 12 AKYVK--GIPLIKYFAEALGPLSEFAWPDLLISTYPKSGTTWVSEILDLIYQEGDLEK 69
QY 77 IGLMNIDQLPVLEYQPQ---GLDIIKELTSPRLIKSHLPVRFPLPSDLHNGDSKVIYMA 132
Db 70 CORAPVFLVRVFLPFSAPGVPTGVLELLKDTTPAPRLKTHLPALLPTLLDQKVKVIYA 129
QY 133 RNPKDLVVSYYQPHRSRLTMSYRGTFQBFRCRRFMNDKLGYSWFEHVOEFWEHRMDSNVL 192
Db 130 RNAKDVAVSYYHYFMAKVVHPDPTGWSFLEKFMAGEVCYGSWYQHQVQEWELSHTHPVL 189
QY 193 FLKYEDMRDLVTWVEQLARFLGVSQCDKAQLEAL---TEHCHQLVDQCCNAELPVG--- 246
Db 190 YLFYEDIKEDPKRETKIAKLEKDISEVLNKIYHTSFDMVKENPMANYTTLFPSSIMDSHI 249
QY 247 -----RGRVGLWKDIFTVSMNEKFDLVYKQKM 273
Db 250 HSISAFMRKGITGDMKSTFTVAQNELFEAHYAKKM 284

RESULT 27
A44011
adrenocortical estrogen sulfotransferase - guinea pig
C:Species: Cavia porcellus (guinea pig)
C>Date: 27-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: A44011
R:Oeda, T.; Lee, Y.C.; Driscoll, W.J.; Chen, H.C.; Strott, C.A.
Mol. Endocrinol. 6, 1216-1226, 1992
A>Title: Molecular cloning and expression of a full-length complementary DNA encoding th
A:Reference number: A44011; MUID:93024479; PMID:1406700
A:Accession: A44011
A>Status: preliminary
A:Molecule type: mRNA; protein
A:Residues: 1-296 <OED>
A:Cross-references: UNIPROT:P49887; UNIPARC:UPI000013622A; GB:U09552; GB:S45979; NID:g494
A:Note: sequence extracted from NCBI backbone (NCBIN:115159, NCBIP:115160)
C:Superfamily: alcohol sulfotransferase
```


Best Local Similarity 32.5%; Pred. No. 5.9e-23; Mismatches 54; Indels 23; Gaps 7;
Matches 89; Conservative 54; Mismatches 108; Indels 23; Gaps 7;
QY 19 YFEFHGVRLPPFCRGK--MEEIAN-PPVRPSDVWIVTPYKSGTSLLOEVVYLVSGADPD 75
Db 4 YTWFGIPFHAGISKETLQNCNFKVKEEDLILAYPKSGTNWLIIEIVCLIQTKGDPK 63
QY 76 EIGLNMIDEQLPVLEYPQGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNP 135
Db 64 WIQSVTIWDRSPWIE--TDVGYDILIKKGPRLMTSHLPMHLFSKSLFSSKAKVILIRNP 122
QY 136 KDLVSVYYQFHRSLTMSYR--LTQFRCRRFMNDKLGYSWFEHVFQEFWEHRMDSNVL 192
Db 123 RDLVSGYGFYFNGNSTLVKKPDSLSGTVEW---FLKGNVLYGSWFHRAWLSMRWDNPL 179
QY 193 FLKYEDMHRDLVTWVEQLARFLVSGCDKAQLEALTEHCHQLV---DOCCNAEALP---- 244
Db 180 LLYIEDMKDTMGTKIKKICDFLGKKLEPDLDLVLKYSSFQVMKENDMSNYSLLMKKSTF 239
QY 245 -----VGRGVLWKDIFTVSMNEKFDLVYKQKM 273
Db 240 TGIGLMRKGTIGDWNHFTVSOAEAFDKVFOEKM 273
RESULT 38
A33569
alcohol sulfotransferase (EC 2.8.2.2) - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 09-Mar-1990 #sequence_revision 09-Mar-1990 #text_change 09-Jul-2004
C:Accession: A33569
R;Ogura, K.; Kajita, J.; Narihata, H.; Watabe, T.; Ozawa, S.; Nagata, K.; Yamazoe, Y.; K
Biochem. Biophys. Res. Commun. 165, 168-174, 1989
A:Title: Cloning and sequence analysis of a rat liver cDNA encoding hydroxysteroid sulfo
A:Reference number: A33569; MUID:90073705; PMID:2590219
A:Accession: A33569
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-284 <OGU>
A:Cross-references: UNIPROT:P15709; UNIPARC:UPI00001709B6; GB:M31363; NID:G204670; PIDN:
C:Superfamily: alcohol sulfotransferase
C:Keywords: sulfotransferase
Query Match 23.0%; Score 351.5; DB 2; Length 284;
Best Local Similarity 32.7%; Pred. No. 8.7e-23;
Matches 90; Conservative 55; Mismatches 105; Indels 25; Gaps 8;
QY 19 YFEFHGVRLPPFCRGK--MEEIAN-PPVRPSDVWIVTPYKSGTSLLOEVVYLVSGADPD 75
Db 4 YTWFGIPFHAGISKETLQNCNFKVKEEDLILAYPKSGTNWLIIEIVCLIQTKGDPK 63
QY 76 EIGLNMIDEQLPVLEYPQGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNP 135
Db 64 WIQSVTIWDRSPWIE--TDVGYDILIKKGPRLMTSHLPMHLFSKSLFSSKAKVILIRNP 122
QY 136 KDLVSVYYQFHRSLTMSYR----GTQFRCRRFMNDKLGYSWFEHVFQEFWEHRMDSNV 191
Db 123 RDLVSGYGFYFNGNSTLVKKPDSLSGTVEW---FLKGNVLYGSWFHRAWLSMRWDNPL 178
QY 192 LFLKYEDMHRDLVTWVEQLARFLVSGCDKAQLEALTEHCHQLV---DOCCNAEALP---- 244
Db 179 LLLYIEDMKDTMGTKIKKICDFLGKKLEPDLDLVLKYSSFQVMKENDMSNYSLLMKKSI 238
QY 245 -----VGRGVLWKDIFTVSMNEKFDLVYKQKM 273
Db 239 FTGTGLMRKGTIGDWNHFTVSOAEAFDKVFOEKM 273
RESULT 39
A26136
senescence marker protein 2, hepatic - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 05-Oct-1988 #sequence_revision 05-Oct-1988 #text_change 09-Jul-2004
C:Accession: A26136

R;Chatterjee, B.; Majumdar, D.; Ozbilen, O.; Murty, C.V.R.; Roy, A.K.
J. Biol. Chem. 262, 822-825, 1987
A:Title: Molecular cloning and characterization of cDNA for androgen-repressible rat liv
A:Accession number: A26136; MUID:87109186; PMID:3805009
A:Accession: A26136
A:Molecule type: mRNA
A:Residues: 1-282 <CHA>
A:Cross-references: UNIPROT:P07631; UNIPARC:UPI000017581B
C:Superfamily: alcohol sulfotransferase
C:Keywords: liver
Query Match 22.1%; Score 338; DB 2; Length 282;
Best Local Similarity 32.0%; Pred. No. 1.3e-21;
Matches 89; Conservative 47; Mismatches 112; Indels 30; Gaps 6;
QY 17 SKYFEFHGVRLP--PFCRGKMEEIAN-PPVRPSDVWIVTPYKSGTSLLOEVVYLVSGAD 73
Db 3 SDYNWFGIPEIPATSYQREILEDIPNKFKVKEEDLILAYPKSGTNWLIIEIVCLIQTKGD 62
QY 74 PDEI-----GLNMIDEQLPVLEYPQGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKV 128
Db 63 PKWIQSCFPGTVPYDEIEWIFRNNHGG-----PRLITSHLPIHLFSKSFSSKAKA 113
QY 129 IYMARNPKDLVSVYYQFHRSLTMSYRGTFOFCRRFMNDKLGYSWFEHVFQEFWEHRMD 188
Db 114 IYLMRNPRIILVSGYFFWGNLTNVLKXNPGSLGTIFYEFWFLQGNVLYGSWFHRAWLSMRWD 173
QY 189 SNVLFLKYEDMHRDLVTWVEQLARFLVSGCDKAQLEALTEHCHQLV---DOCCNAEALP---- 236
Db 174 DNFLVLYIEDMKDTMGTKIKKICDFLGKKLEPDLDLVLKYSSFQVMKENDMSNYSLIKE 233
QY 237 CCNAEALPVGR-GRVGLWKDIFTVSMNEKFDLVYKQKM 273
Db 234 DPILTGLMRKGTIGDWNHFTVSOAEAFDKVFOEKM 271
RESULT 40
B40216
flavonol 3'-sulfotransferase - Flaveria chloraefolia
C:Species: Flaveria chloraefolia
C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: B40216
R;Varin, L.; Deluca, V.; Ibrahim, R.K.; Brissson, N.
Proc. Natl. Acad. Sci. U.S.A. 89, 1286-1290, 1992
A:Title: Molecular characterization of two plant flavonol sulfotransferases.
A:Reference number: A40216; MUID:92159034; PMID:1741382
A:Accession: B40216
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-311 <VAR>
A:Cross-references: UNIPROT:P52836; UNIPARC:UPI000012A3FC; GB:M84135; NID:G168166; PID:G
A:Experimental source: terminal bud
A:Note: sequence extracted from NCBI backbone (NCBI:82218)
C:Superfamily: alcohol sulfotransferase
Query Match 19.8%; Score 303; DB 2; Length 311;
Best Local Similarity 28.3%; Pred. No. 1.5e-18;
Matches 82; Conservative 57; Mismatches 95; Indels 56; Gaps 11;
QY 30 FCRGMEEIANPPVRPSDVWIVTPYKSGTSLLOEVVYLVSGADPDEIG-LMNI--DEQL 86
Db 34 FLEGRILSEQFKAHNDVFLASYPKSGTTLKAWICITITREKFDSDTSPLTTMPHDCI 93
QY 87 PVLEYPOQGLDIKE-----LTSPLRIKSHLPYRFLPSDLHNGDSKVIYMARNPKDLVVS 141
Db 94 PLLR---KDEKIQENQNSLYTP--ISTHYYKSLPESARTSNCKVIYIYRNMKDVLVS 148
QY 142 YYQFHRSLTMS-----YRGTFQFCRRFMNDKLGYSWFEHVFQEFWEHRMDSN--VLFL 194
Db 149 YYHFLRQIVKLSVEEAPPEEAFDFCQGISS----CGPYWEHIKGYWKASLEKPEIFLFL 204
QY 195 KYEDMHRDLVTWVEQLARFLG-----VSCDKAQLEAL-----TE 228

Db 217 HVLFMKFEEMKAPRDQIKKFAEFLG--CPFTKEEESGVSDEIIDLCSRLNLSLEINK 274
Qy 247 -----RGRVGLWKDIFTVSMNEKFDLVVYKQKMGKCDLTF 280
Db 275 TGKLNSGRENKMFRRKGEVGDWKNVLTPEMKNKIDWIOEKLQNSGLKP 323

RESULT 44
T07833
probable steroid sulfotransferase (EC 2.8.2.15) 3 - rape
C:Species: Brassica napus (rape)
C>Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 09-Jul-2004
C:Accession: T07833
R:Richard, M.; Nicolle, L.; Varin, L.
submitted to the EMBL Data Library, April 1997
A:Reference number: Z16161
A:Accession: T07833
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-325 <RIC>
A:Cross-references: UNIPROT:O82410; UNIPARC:UPI00000AA137; EMBL:AF000307; NID:G3420007;
C:Genetics:
A:Gene: ST3
C:Superfamily: alcohol sulfotransferase
C:Keywords: sulfotransferase

Query Match 18.3%; Score 280; DB 2; Length 325;
Best Local Similarity 26.6%; Pred. No. 1.6e-16;
Matches 82; Conservative 60; Mismatches 118; Indels 48; Gaps 12;

Qy 9 PSTGGEFSKYFEFHG-VRLPFCRGKMBEIANFFVRPSDVMIVTPKSGTSLLOEVVYL 67
Db 28 PSEKGLVSIQYQGRWHTALLQGLTCQKHAKDSDIILVTNPKSGTTLWLSLVA 87
Qy 68 -----VSQADPDEIGLWIDQLPVLE---YPOGLDIIKELTSPRLIKSHLPYRF 116
Db 88 LINRHKFPVSSGDHP--LLVTNPHLLVPFMGYYESED--FSLPPLPRLNTHSHLS 144

Qy 117 LPSDLHNGDSKVIYMARPKDLVSYVQFHRSL---RTMSYRGTFQFCRRFRMNDKLGVG 173
Db 145 LPESVSKSSCOIYVCRPNKMFVLSLWFGKKLAPQETADY--PLEKAVEACQCKFTAG 202

Qy 174 SWFEHVQFWEHRMD--SNVLFLKYEDMHRDLVTWVEQLARFLGVSCDKAQLEALTEHCH 231
Db 203 PFWDHVLEWYASLEPNKVLFTVTEELKKQFEVEVGRIAEFIG--CGTAESEVSE--- 257

Qy 232 QLVDOCCNAEA-----LPVG-----RGRVLWKDIFTVSMNEKFDLVVYKQK 272
Db 258 --IVKLCSFESLSRLEVNROGKLPNGIETNAFFRKGEIGGWRDTLSLADAIDRTTEK 315

Qy 273 MGKCDLTF 280
Db 316 FGGSGLKPF 323

RESULT 45
E86319
probable flavonol sulfotransferase [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: E86319
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Dewar, T.H.; Dewar, K.;
ansen, N.P.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Frazer, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: AB6141; MUID:21016719; PMID:11130712

A:Accession: E86319
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-346 <STO>
A:Cross-references: UNIPROT:Q9FZ80; UNIPARC:UPI000000C631; GB:AE005172; NID:G9795597; P1
C:Genetics:
A:Map position: 1
C:Superfamily: alcohol sulfotransferase

Query Match 18.2%; Score 278.5; DB 2; Length 346;
Best Local Similarity 24.9%; Pred. No. 2.3e-16;
Matches 83; Conservative 62; Mismatches 119; Indels 69; Gaps 12;

Qy 1 MAESEAE-TPSTPGGEF--SKYFEFHGVRLP-----PFCRGKME 36
Db 11 VSESHNELASSSPSEFEKNQKHQYQBIATLPHKDGWRPKDPVEYGGHWWLQPLLEGLIH 70

Qy 37 EIANFPVPSDVMIVTPKSGTSLLOEVVYLSQCADPD-----EIGLWIDQLPVLE-- 90
Db 71 AQKFFKARPNDFFVCSYPKTGTTLKALTFAIANRSKFDVSTNPLKRNPHFEFVPIEID 130

Qy 91 YP-QPGLDIIKELTSPRLIKSHLPYRFPLPSDLHNGDSKVIYMARPKDLVSYVQFHRSL 149
Db 131 PFPFSDVVLKD-EGNTLFTSTHPIYDLLPESVVGCKIVYIWRDPKDTFVSMWTFAPKE 189

Qy 150 RT-----MSYRGTFQFCRRFRMNDKLGYSWFEHVQEFWE--HRMDSNVLFYKEDMHRD 202
Db 190 RSQGPVSVSIEEAFDKYCQGLS----AYGPYLDHVLGYWKAYQANPDQILFLKYETMRAD 245

Qy 203 LVTWVEQLARFLGVSCDKAQLEALTEHCHQLVDOCCNAEALP----- 244
Db 246 PLPYVKRLAEFMGYGFTKEEBEG---NVVEKVKVLCSPFETLKNLEANKGEKDRDPVAVY 302

Qy 245 -----VGRGRVLWKDIFTVSMNEKFDLVVYKQK 272
Db 303 ANSAVFRKGKVDWQNYLTPEKVARIDGLMEEK 335

RESULT 46
A84452
probable steroid sulfotransferase [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: A84452
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: A84452
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-326 <STO>
A:Cross-references: UNIPROT:P52839; UNIPARC:UPI00000484E4; GB:AE002093; NID:G4406767; P1
C:Genetics:
A:Gene: At2G03760
A:Map position: 2
C:Superfamily: alcohol sulfotransferase

Query Match 18.1%; Score 277.5; DB 2; Length 326;
Best Local Similarity 26.5%; Pred. No. 2.6e-16;
Matches 82; Conservative 57; Mismatches 121; Indels 49; Gaps 12;

Qy 9 PSTGGEFSKYFEFHG-VRLPFCRGKMBEIANFFVRPSDVMIVTPKSGTSLLOEVVYL 67
Db 28 PKKXGLVSEIYFQGLMHTQAILQGLICQKFEAKDSDIILVTNPKSGTTLWLSLVA 87

Qy 68 V-----SQGADPDEIGLWIDQLPVLE---YPOGLDIIKELTSPRLIKSHLPYR 115
Db 88 LLNRHKFPVSSGNHP--LLVTNPHLLVPFLEGVYYESPDFD--FSSLSFRLMNTSHSL 144

Qy 116 FLPSDLHNGDSKVIYMARPKDLVSYVQFHRSL---RTMSYRGTFQFCRRFRMNDKLG 172

Db 145 SLPSVSSKSIIVCCRNPKDMFVSLWHFGKLAPEADY--PIEKAVAFCEGKFIG 202
Qy 173 GSWFEHVQEFW--EHRMDSNVLFLKYEDMHRDLVTMVEQLARFLGVSCDKAQLAETHC 230
Db 203 GPFWDHILEYVWYASRENPNKLVFVYBELKKQTEVEMKRIAEFL--ECGFIEEBEVRE-- 258
Qy 231 HOLVDQCCNAA-----LPVG-----RGRVGLWKDIFTVSMNEKFDLVYKQ 271
Db 259 ---IVKLCSPESLSNLEVNKEGLPNGIETKTFPRKGBIGGWRDRTLSESLABEIDRTIEE 315
Qy 272 KMGKCDLTF 280
Db 316 KFKGSLKF 324

RESULT 47

T07831

probable steroid sulfotransferase (EC 2.8.2.15) 1 - rape

C:Species: Brassica napus (rape)

C:Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 09-Jul-2004

C:Accession: T07831

R:Richard, M.; Nicolle, L.; Varin, L.

submitted to the EMBL Data Library, April 1997

A:Reference number: Z16161

A:Accession: T07831

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-323 <RIC>

A:Cross-references: UNIPROT:O82408; UNIPARC:UPI00000A1423; EMBL:AF000305; NID:g3420003;

C:Genetics:

A:Gene: ST1

C:Superfamily: alcohol sulfotransferase

C:Keywords: sulfotransferase

Query Match 18.1%; Score 276.5; DB 2; Length 323;
Best Local Similarity 27.4%; Pred. No. 3.1e-16;
Matches 85; Conservative 59; Mismatches 113; Indels 53; Gaps 14;
Qy 9 PSTGFEFSKYFPHGV-RLPFCRGKMBEIANFVRPSDWMIVTPYKSGTSLLOEVVYL 67
Db 27 PSEKWLVSQMYQFEGIQTOALVQGVNVCQKHFANDSDVILATLAKSGTTWLKALLFA 86
Qy 68 V-----SQGADPDEIGLMDIDQLPVE---YFQPGLDIIKELTSRLIKSHLPYRFL 117
Db 87 LIHRHKFPVSGKHP--LLVTNPHSLVPYLEGDCYSSPEVN-FAELPSRLMQTLTHHSL 143
Qy 118 PSDLHNGDSKVIYMARNPKDLVSVYQFHRSL---RTMSY--RGTFQFCR-RFMNDKLG 171
Db 144 PVSISKSSCKIYCCRNPKDMFVSIWHFGKLAPEKTAEPYPIETAVAAFCCKGKFIG--- 199
Qy 172 YGSMFEHVQEFWEHRM--DSNVLFKLYEDMHRDLVTMVEQLARFLGVSCDKAQLAETHC 229
Db 200 -GPFWDHILEYVWYASRENPNKLVFVYBELKKQTEVEMKRIAEFTG--CGFTAEEBESVSE- 255
Qy 230 CHQLVDQCCNAA-----LPVG-----RGRVGLWKDIFTVSMNEKFDLVYK 270
Db 256 ---IVKLCSPESLSNLEVNROGKLPNGTIESNAFFRKGTGGRDRTLSESLADVIDRTTE 311
Qy 271 QMGKCDLTF 280
Db 312 QKFGSGLKF 321

RESULT 48

A40216

flavonol 4'-sulfotransferase - Flaveria chloraefolia

C:Species: Flaveria chloraefolia

C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 20-Sep-1999

C:Accession: A40216

R:Varin, L.; Deluca, V.; Ibrahim, R.K.; Brisson, N.

Proc. Natl. Acad. Sci. U.S.A. 89, 1286-1290, 1992

A:Title: Molecular characterization of two plant flavonol sulfotransferases.

A:Reference number: A40216; MUID:92159034; PMID:1741382

A:Accession: A40216

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-320 <VAR>

A:Cross-references: UNIPARC:UPI0000175837

A:Experimental source: terminal bud

A:Note: sequence inconsistent with the nucleotide translation

A:Note: sequence extracted from NCBI backbone (NCBIN:82216, NCBI:82217)

C:Superfamily: alcohol sulfotransferase

Query Match 18.0%; Score 275.5; DB 2; Length 320;
Best Local Similarity 27.5%; Pred. No. 3.8e-16;
Matches 80; Conservative 63; Mismatches 111; Indels 37; Gaps 11;
Qy 18 KYFHFHGVRLPFCRGKMBEIANFVRPSDWMIVTPYKSGTSLLOEVVYVLSQGADPDEI 77
Db 34 KYQDFWG--LQNNIEGAILAQOSFKARPDVFLCSYPSKSGTTWLKALAYAVTREKPFDEF 91
Qy 78 G---LMNIDEQ-LPVLVYQPGLDIIKELTSR-----LIKSHLPYRFLPSDLHNGDSKV 128
Db 92 TSPLLTNIPHNCPYIE-----KOLKKTVDNNNSCFTPMATHMPYHYVLPKSLIALNCKM 146
Qy 129 IYMARNPKDLVSVYQFHRSLRTMSY-RTQEFRCRRFMDKLGYSGFHVFQFWEHRM 187
Db 147 VYIYRNKDVIVSFYHFGREITKPLEDAPFEAEFDEYHGISQFPGYMDHLGLGYKASL 206
Qy 188 D--SNVLFKLYEDMHRDLVTMVEQLARFLGVSC-----DKAQLAETHC-----HOLV 234
Db 207 ERPEVILFLKYEDVKDPTSNVKLAELFYGYPFTFEEKEGVIESIIKLCSPENLSNLEV 266
Qy 235 DQCCNAEA-LPV-----GGRGVGLWKDIFTVSMNEKFDLVYKQMGKCDL 278
Db 267 NKGNSRGFLPIENLYFRKAKDGDWKNYFTDEMTEKIDKLIDEKLSATGL 317

RESULT 49

H96768

protein flavonol sulfotransferase F2P9.4 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C:Accession: H96768

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

ansen, N.F.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

Nathure 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: H96768

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-350 <STO>

A:Cross-references: UNIPROT:Q9C9C9; UNIPARC:UPI000000C528; GB:AE005173; NID:g7109463; PI:

C:Genetics:

A:Gene: F2P9.4

A:Map position: 1

C:Superfamily: alcohol sulfotransferase

Query Match 17.5%; Score 267; DB 2; Length 350;
Best Local Similarity 25.5%; Pred. No. 2.3e-15;
Matches 83; Conservative 56; Mismatches 127; Indels 60; Gaps 12;
Qy 2 AESEATSTGCESESKYF-----EFHGV-RLPFCRGKMBEIAN 40
Db 19 SHDETFTSTEFKNQRYQDLSTFPPEKGNRPKEPLIEYGGVWNLPSLEGCGTHAQEF 78
Qy 41 FVVRPSDWMIVTPYKSGTSLLOEVVYVLSQGADPDE----ICLMNIDQLPVLE--YP-Q 93

Db 79 FOARPSDFLVCSYPKTGTWTKALTAFAIANRGRFDDSSNPLKRNPHETVPVIEIDFPFF 138

Qy 94 PGLDIIKELTSRLIKSHLPYRFLPSDLHNGDSKVIYMARNPKDLVSVYQFHRSLRT-- 151

Db 139 PEVDVLKD-KGNTLFSHTHPIYELLPPDSVVKSGCKMYIWRPKDFTISMWTFLLHKERTEL 197

Qy 152 ---MSYRGTFQBFRCRPMNDKLGYSWFHVEQFWEHRMDS--NVLFLKYEDMHRDLVTM 206

Db 198 GPNVLEESFDMFCRGLS-----GYGFLNHLAYWKAYQENPDRILFLKYETMRADPLPY 253

Qy 207 VEOLARFLGVSC-----DKAOLEALTEHC--HQLVDCCN-----ABALPVG 246

Db 254 VKSLAEFMHGHTABEEKGVEKVVNLCSTFETLKNLEANKGEKDEDRPGVYANSAYPR 313

Qy 247 RGRVGLWKDIFTVSMNEKFDLVYKQK 272

Db 314 KGKVGDSNYLTPEMAARIDGLMEEK 339

RESULT 50

T07832

probable steroid sulfotransferase (EC 2.8.2.15) 2 - rape

C:Species: Brassica napus (rape)

C:Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 09-Jul-2004

C:Accession: T07832

R:Richard, M.; Nicolle, L.; Varin, L.

submitted to the EMBL Data Library, April 1997

A:Reference number: Z16161

A:Accession: T07832

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-324 <R1C>

A:Cross-references: UNIPROT:O82409; UNIPARC:UPI0000046BEA; EMBL:AF000306; NID:G3420005;

C:Genetics:

A:Gene: ST2

C:Superfamily: alcohol sulfotransferase

C:Keywords: sulfotransferase

Query Match 17.4%; Score 265.5; DB 2; Length 324;

Best Local Similarity 26.8%; Pred. No. 2.8e-15;

Matches 80; Conservative 61; Mismatches 128; Indels 29; Gaps 11;

Qy 9 PSTPGEFSKYFEFHG-VRLPPFCRGKMEIEANFPVRPSDVWIVYPKSGTSLLOEVY- 66

Db 28 PSEKGLVQCMYQFOGRWHTQALLOGILTCQKHFEAKSDIILVTNPKSGTTWLKALVFA 87

Qy 67 LVSGADPDEIGLWNIDEQLPVLEYQPGLDIIK-----ELTSPRLIKSHLPYRFLPS 119

Db 88 LINRHKFP-VYSVIIISCYQSALLVPFLGRSLRSPDFPSQUSSPRLNTHISHLSLPE 146

Qy 120 DLHNGDSKVIYMARNPKDLVSVYQFHRSL---RTMSYRGTFQBFRCRPMNDKLGYSWF 176

Db 147 SVKSSCKIYVCCRNPKDMFVSLWHFGKXLAPEETADY--PIEKAVEAFCQCKFIGGPPW 204

Qy 177 EHVQFWEHRMD--SNVLFLKYEDMHRDLVTMVEQLARFLGVS-CDKAOLEALTEHC--H 231

Db 205 DHVLEYWYASLENPNKVLFSVSEEPKKTGETIKRIAEFLGGLGVGEESVRAIVKLCSPE 264

Qy 232 QLVDQCCNAA-LPVG-----RGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280

Db 265 SUSSLEVNREGKLPDSGMETRAFRKGEVGGWRDITLTSUAEVIDRTIEEKFOGSLKF 322

RESULT 51

F86407

probable sulfotransferase F3H9.17 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C:Accession: F86407

R:Theologis, A.; Ecker, J.R.; PalM, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Mafti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: F86407

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-326 <STO>

A:Cross-references: UNIPROT:O9F291; UNIPARC:UPI00000482DF; GB:AE005172; NID:g9795621; P1

C:Genetics:

A:Map position: 1

C:Superfamily: alcohol sulfotransferase

Query Match 17.2%; Score 262.5; DB 2; Length 326;

Best Local Similarity 27.2%; Pred. No. 5.2e-15;

Matches 85; Conservative 56; Mismatches 112; Indels 59; Gaps 13;

Qy 14 PFES-KYFEFHGVRLLPPFC-----RGKMEIEANFPVRPSDVWIVYPKSGTSLLOEVV 65

Db 29 DFDSTKLFKYQG-----CWYDDKTQGVNLFQRFEPQDTIIIASFPKSGTTWLKALT 82

Qy 66 YLV-----SQGADPDEIGLWNIDEQLPVLEY-----PQGLDIIKELTSRLIKSHLP 113

Db 83 VALLERSKQKHSSDDHLLDNPHGLVPLFLRLFTSTSKPDLTSTIS--SSPRLFTSTHVA 140

Qy 114 YRFLPSDLHNGDSKVIYMARNPKDLVSVYQFHRSLRTMSYR----GTQFQFCRRFMND 168

Db 141 FQTLREALNKSPCKIYVVRNVKDLVSVFWFNSAKLIEBERSILDSMFESFCRGVIN- 199

Qy 169 KLGYSWFHVEQFWEHRMD--SNVLFLKYEDMHRDLVTMVEQLARFLGVSCKAQLAEL 226

Db 200 ---YGPSMEHVLNLYWRASLEDSKNVFLKYBELKEPRVQLKRLAEFL--DCPFTVEEBE 254

Qy 227 TEHQQLVDQCC-----NAEALPVG-----RGRVGLWKDIFTVSMNEKFDLV 268

Db 255 RGSVEEILDCLSLRNKLNKTKTLRGADHKIFPRKGEVGDGSKNHLTPEMEKIIDMI 314

Qy 269 YKQKMGKCDLTF 280

Db 315 TEEKEGSDLKF 326

RESULT 52

S69188

probable flavonol sulfotransferase (EC 2.8.2.-) - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 15-Nov-1996 #sequence_revision 27-Feb-1997 #text_change 21-Jul-2000

C:Accession: S69188

R:Lacomme, C.; Roby, D.

Plant Mol. Biol. 30, 995-1008, 1996

A:Title: Molecular cloning of a sulfotransferase in Arabidopsis thaliana and regulation

A:Reference number: S69188; MUID:96270377; PMID:8639757

A:Accession: S69188

A:Molecule type: mRNA

A:Residues: 1-302 <LAC>

A:Cross-references: UNIPARC:UPI0000046BEB; EMBL:Z46823; NID:g599639; PIDN:CAA86850.1; P1

C:Superfamily: alcohol sulfotransferase

C:Keywords: sulfotransferase

Query Match 17.0%; Score 259.5; DB 2; Length 302;

Best Local Similarity 26.9%; Pred. No. 8.5e-15;

Matches 77; Conservative 52; Mismatches 108; Indels 49; Gaps 12;

Qy 9 PSTPGEFSKYFEFHGV-RLPPFCRGKMEIEANFPVRPSDVWIVYPKSGTSLLOEVYL 67

Db 28 PREKGLVSEIYEFQGLWHTQAILQGLICQKRFKAKSDIILVTNPKSGTTWLKALVFA 87

Qy 68 V-----SOGADPDEIGLWNIDEQLPVLE---YPOGLDIIKELTSRLIKSHLPVR 115

Db 88 LLNRHKFPVSSSGNHP--LLVTNPHLLVPFLFEGVYESPDFD-FSSLSPLRNTHSHL 144

Db 77 SGTTLWKALTFALVQSRKSHLEDHQHPLLHNPHEI-VPNLELDL-YLKSSKP--DLTKP 132

Qy 102 L-----TSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPKDLVVSYYQFHRSLR-----TM 152

Db 133 LSSSSSSPRLFSFTHMSLDLPQVLKENLCKIVVVCNVDVMVSVVMYFQSKKXITRAEDY 192

Qy 153 SYRGTFQFCRRFMDNKLGYSGWFEHVQFWEHRMD--SNVFLKYEDMHRDLVTMVEOL 210

Db 193 SLEALFESFC----NGVTLHGFPWDHALSYWRGSLDDPKHFLFMRVEDLKASBPTQVKRL 248

Qy 211 ARFLGVSC-----DKAQLEALTEHCHQLVDQCCNAEALPVG-----RG 248

Db 249 AEFL--DCPFTKEEEDSGSVDKLELC-----SLSNLRSVEINKTTSRVDKSFYRKG 301

Qy 249 RVLGWLKDIFTVSMNEKFDLVYKQMGKCDLTF 280

Db 302 QVGDWKSMTPEMVDKIDMIIBEKLGSLGF 333

RESULT 56

T06012

hypothetical protein T25K17.90 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C>Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jul-2004

C:Accession: T06012

R:Devan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X submitted to the Protein Sequence Database, March 1999

A:Reference number: Z15382

A:Accession: T06012

A:Molecule type: DNA

A:Residues: 1-314 <REV>

A:Cross-references: UNIPROT:Q9ST06; UNIPARC:UPI00000489F1; EMBL:AL049171; GSPDB:GN00062;

A:Experimental source: cultivar Columbia; BAC clone T25K17

C:Genetics:

A:Gene: ATSP:T25K17.90

A:Map position: 4

A:Introns: 109/3

C:Superfamily: alcohol sulfotransferase

Query Match 15.4%; Score 235.5; DB 2; Length 314;

Best Local Similarity 25.6%; Pred. No. 1.1e-12;

Matches 69; Conservative 58; Mismatches 104; Indels 39; Gaps 10;

Qy 30 FCRGMEEIANPVRPSDVWIVTPKSGTSLLOEV-VYLVSGADPDEIGLNMIDQL-- 86

Db 46 FLOGVLNFORGFKPQDTHIVASPKSGTLWLKALTVALFERTKNPSH-----DPMASH 99

Qy 87 PVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPKDLVVSYYQFH 146

Db 100 PLLSNPNHNL-----LSSSSPRLFSFTHPTTLQAVKDSCKVYVICRDAKSLVS--RWH 154

Qy 147 ---RSLRTMSYRGTFQFCRRFMDNKLGYSGWFEHVQFWEHRMD--SNVFLKYEDMHR 201

Db 155 IVCRSLNKEEDRTILESMPESFCSGVCLFGPFDWHLISYWKASLEKPKQVLFMRVDEIKT 214

Qy 202 DLVTMVEQLARFLGVSCDKAQLEALTEHCHQLVDQCC--NABALPVG----- 246

Db 215 DPHGQLKLAELFLG--CPSPKEEKNGSLNKLILEMCSLPNLSLEVNKTKGKSINGIEVKN 272

Qy 247 ---RGVGLWKDIFTVSMNEKFDLVYKQMGK 273

Db 273 HPRKRGVGDWKNHLTPEMGSKIDMIMKEKL 302

RESULT 57

H84451

probable steroid sulfotransferase [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004

C:Accession: H84451

R:Lin, X.; Kaul, S.; Rounslev, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: H84451

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-331 <STO>

A:Cross-references: UNIPROT:Q9ZPQ6; UNIPARC:UPI00000484E3; GB:AE002093; NID:G4406766; P

C:Genetics:

A:Gene: At2G03750

A:Map position: 2

C:Superfamily: alcohol sulfotransferase

Query Match 14.7%; Score 224.5; DB 2; Length 331;

Best Local Similarity 25.1%; Pred. No. 1e-11;

Matches 68; Conservative 53; Mismatches 117; Indels 33; Gaps 8;

Qy 40 NPPVRPSDVWIVTPKSGTSLLOEVVYLVSQ-----GADPDEIGLNMIDQLPVLE--- 90

Db 63 HFKPRDTHIIILASLPRGGTTLKSLIFAVVHREKVRGTPTQTHPLLQNPDLVFPFLEVEL 122

Qy 91 YPQGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPKDLVVSYYQFHRSL- 149

Db 123 VANSQIPDLAKYSSPMIFSTHMLQAL-REATTKACTVYVCRGKIDTFVSGWHYRNMLH 181

Qy 150 RTMSYRGTFQFCRRFMDNKLGYSGWFEHVQFWEHRMDS--NVFLKYEDMHRDLVTMV 207

Db 182 RTKMDQATFELMFDAYCRGVLLYGPYEHVLSYWKSGLEAKENVLFMKYEIIIEPRVQV 241

Qy 208 EQLARFLGVSCDKAQLEALTEHCHQLVDQCC--NAEALPVG-----RGR 249

Db 242 KRLAEFL--ECPFTKEEESGVBEILKLCSLRNLSNLEVNKNGTTRIGVDQSVQFPRKGE 299

Qy 250 VGLWKDIFTVSMNEKFDLVYKQMGKCDLTF 280

Db 300 VGDWKNHLTPQMAKTFDEIIDYRLGDSGLIF 330

RESULT 58

A95936

probable alcohol sulfotransferase (EC 2.8.2.2) [imported] - Sinorhizobium meliloti (stra

C:Species: Sinorhizobium meliloti

C>Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004

C:Accession: A95936

R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan

Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001

A:Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo

A:Reference number: A95842; MUID:21396508; PMID:11481431

A:Accession: A95936

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-280 <KUR>

A:Cross-references: UNIPROT:Q92VF3; UNIPARC:UPI000000CB647; GB:AL591985; PIDN:CAC49153.1.

A:Experimental source: strain 1021, megaplasmid pSymB

R:Galbert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,

pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;

L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001

A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,

hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.

A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.

A:Reference number: A96039; MUID:21368234; PMID:11474104

A:Contents: annotation

C:Genetics:

A:Gene: Smb21249

A:Genome: plasmid

C:Keywords: sulfotransferase

Query Match 10.5%; Score 161; DB 2; Length 280;

Best Local Similarity 23.9%; Pred. No. 2.5e-06;

Matches 64; Conservative 46; Mismatches 98; Indels 60; Gaps 12;

Qy 46 SDWIVTPKSGTSLLOEVV--YLVSGADPDEIGLNMIDQLPVLEYPQGLDIKELT 103

Db 25 ADSFLISPKSGRTWFRVLSHYLATIARVETIDLHMFSIVNFD-----LDVRGNP 79
Qy 104 S-----PRLIKSHLPYR---FLPSDLHNGDSKVIYARNPKDLVSVYQFPHRSR 150
Db 80 GVRPREAKDAIPTILVSHLDYRASLFL-----RRPVIMVVRDPRDIVSAY-PHATRH 131
Qy 151 TMSYRGTFQFCRRPMNDKLVGSGWFHVFQFWEHRMDSNVLF-IKYEDMRDLVTWVEQ 209
Db 132 KHRFAGTLTEFIK---DRDQGMKIDYING-WAAGLSNRAHFVLSYEGLSADTDGRTEA 187
Qy 210 LARFLVGSQCDKAQLEALTE-----HCHQLVDQCCNAEALPVGRGRVG 251
Db 188 VLEFLURCPVDRTRALRAAVAGFEAMQDRVERVEGIPAHDYRND-----VESLRMRGKAG 243
Qy 252 LMKDIFTVSMNEKFDLVYKQKMGKCDLT 279
Db 244 GFRDYL-----DEAQVAEVLCAADLT 266
RESULT 59
E84072
hypothetical protein BH3381 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C:Accession: E84072
R:Takami, H.; Nakase, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: E84072
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-249 <STO>
A:Cross-references: UNIPROT:Q9K710; UNIPARC:UPI00000C41A8; GB:AP001518; GB:BA000004; NID
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH3381

Query Match 7.2%; Score 110; DB 2; Length 249;
Best Local Similarity 21.3%; Pred. No. 0.054;
Matches 56; Conservative 54; Mismatches 101; Indels 52; Gaps 11;
Qy 44 RPSDWIIVTYPKSGTSLQLQEVVLYVSQADPDEIGLMDIDOLPVLEVPQGLDIKELT 103
Db 4 KASKVLITSVPKCGTHLLKQVYL-----GAPFYKRG-----KWISLRESIREQIDNCHDQT 54
Qy 104 SPRLIKSHLPY-----RFLPSDLHNGDSKVIYARNPKDLVSVYQFPHRSRLT----- 151
Db 55 ---IMVGHLEYSEEIVRY-----LNEKGIRRLFIYDRDMLISYFYVYKKEIAPKHELNY 107
Qy 152 -MSYRGTFQFCRR-----FMNDKLVGSGWFHVFQFWEHRMDSNVLFKYEDM---HR 201
Db 108 FLHVLGSDEERIMKLTITGPHYQEENLTGDLTLHYQPLPWTIEDKGTALRFDLVSSNE 167
Qy 202 DLVTWVEQLARL--GVSCDKAQLAEALTEHCHQLVDQCCNAEALPVGRGRGLMKDITV 259
Db 168 CKEQITISKIADYLMHGSTFDQYNKAVLK--RMLENTRPSHSGTFRKGVSGDWKNYFTA 224
Qy 260 SMNEKF-----DLVYKQKM 273
Db 225 QHKSVPKEVAGOLLIDLGYEKM 247
RESULT 60
G70393
conserved hypothetical protein aq_1088 - Aquifex aeolicus
C:Species: Aquifex aeolicus
C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Oct-2004
C:Accession: G70393
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov
V.
Nature 392, 353-358, 1998

A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MUID:98196666; PMID:9537320
A:Accession: G70393
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-761 <AOF>
A:Cross-references: UNIPROT:O67178; UNIPARC:UPI0000056515; GB:AE000722; NID:G2983559; PI
A:Experimental source: strain VP5
C:Genetics:
A:Gene: aq_1088
F:103-136/Domain: tetratricopeptide repeat homology <TT1>
F:137-170/Domain: tetratricopeptide repeat homology <TT2>
F:171-203/Domain: tetratricopeptide repeat homology #status atypical <TT3>
F:204-237/Domain: tetratricopeptide repeat homology <TT4>
F:317-350/Domain: tetratricopeptide repeat homology <TT5>
F:351-384/Domain: tetratricopeptide repeat homology <TT6>
Query Match 7.2%; Score 109.5; DB 2; Length 761;
Best Local Similarity 21.5%; Pred. No. 0.27;
Matches 49; Conservative 39; Mismatches 67; Indels 73; Gaps 13;
Qy 44 RPSDWIIVTYPKSGTSLQLQEV-----VYLVSQ-----GADPDE--I 77
Db 488 RP--IFVLGMPRSGTTLTEQILGSHSMVYPAGELPFVVKIVNLIPKALQYVYVKEPKWPE 545
Qy 78 GLMIDIDOLPVLEVPQGLDIKELTS--PRLIKSHLPYR-----LPSDLHNGDS 126
Db 546 AILFEDERL-LKSAQYVLDRAKLDSEHPRIV-DKLPNFDYVGLILLMFP-----NA 597
Qy 127 KVIYARNPKDLVSVYQFPHRSRMTSYRGTFQFCRRFMNDKLVG---SWFEH----- 178
Db 598 KVIHLKRNLDLVAVSNYQ-----QNFPAK--HGTWGFAPDLRWIGHMLNDH 641
Qy 179 --VQSFMEHRMDSNVLFKYEDMRDLVTWVEQLARFLGVSCDKAQL 224
Db 642 RAIMEHWHKLPFGQIYELDYQLTEPEPEVIRELLEFCLELPWEDRVLE 699
RESULT 61
B70958
hypothetical protein Rv1373 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: B70958
R:Coile, S.I.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.;
Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: B70958
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-326 <COL>
A:Cross-references: UNIPROT:P71801; UNIPARC:UPI00000C14AA; GB:Z81011; GB:AL123456; NID:G
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: Rv1373
Query Match 7.0%; Score 106.5; DB 2; Length 326;
Best Local Similarity 23.7%; Pred. No. 0.16;
Matches 65; Conservative 37; Mismatches 105; Indels 67; Gaps 14;
Qy 43 VRPSDWIIVTYPKSGTSLQLQEVV-YLVSQGAD-PDEIGLMN--IDEQLPVLEYPQGLDI 98
Db 28 LRDGDIIISAPSKSGLTWTQRLVSLVDFDPLDPLGPLSTVSPWLDQTRPIEVEVATLDA 87
Qy 99 IKELTSPRLIKSHLPYRFLPSDLHNGDSKVIY--NARNPKDLVSV-YVQ----- 144
Db 88 QQH---RRFIKTHPTPLDGLVL-----DDRVSYICVGRDPRDAVSMLYQSANNEDRMRI 139
Qy 145 -----FHRSL-----RTMSYRGTFQFCRRFMN--DKLVGSGWFHVFQ----- 180

Db 140 LHEAVVPFHERIAPPFAELGHARSPTTEP-RDWMGPNQPPGIGFTHLKGIGTTLANTILH 198
Qy 181 ----EFWEHRMDSNVFLFKYEDMHRDLVTWVEQLARFLGVSCDK-----AQLEALTE 228
Db 199 QLGTVWVRHLENVALFHYADYQADLAGELLARVGLGTAATRDARDLAQVATLDAMRS 258
Qy 229 HCHQL-----VDCCNAEALPVGRGVLWKDIFT 258
Db 259 RASEIAPNTDGIWHSDFRFRGGSGDWQQPFT 292

RESULT 62
T29144
partial CDS - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
A:Accession: T29144
R:Pauley, A.; Gattung, S.
submitted to the EMBL Data Library, July 1996
A:Description: The sequence of C. elegans cosmid K11C4.
A:Reference number: Z20577
A:Accession: T29144
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-5107 <PAU>
A:Cross-references: UNIPROT:Q94279; UNIPARC:UPI000011013D; EMBL:U64854; PIDN:AAB18318.1;
A:Experimental source: strain Bristol N2; clone K11C4
C:Genetics:
A:Gene: CESP:unc-68
A:Map position: 5
A:Introns: 27/1; 64/3; 92/3; 127/1; 158/2; 1222/2; 1300/2; 1347/2; 1391/1; 1419/3; 1517/3; 3269/2; 3313/2; 3466/1; 3519/3; 3629/3; 3658/2; 3710/1; 3741/3; 3779/2; 3810/3
C:Superfamily: ryanodine receptor; transcription initiation factor sigma region 1 homolo

Query Match 6.6%; Score 101; DB 2; Length 5107;
Best Local Similarity 21.7%; Pred. No. 20;
Matches 79; Conservative 43; Mismatches 138; Indels 104; Gaps 14;
Qy 7 ETPSPGFESKYFFHGVRLPP-----FCRGK-----MEEIANFPVRPSDVMIV 51
Db 1962 QTDMPFAEAAKTKTF---RCPPEQMFRLLMCKYKEERDPELMEEDADVQCPWAEGIQ 2018
Qy 52 TYPKSGTSLLOEVVYLVSGQADPDEIGLWNIDE-----QLPVLEYQPQGLD 97
Db 2019 QQLRDFCELLVKGIGNVKXEGSDDDQLALTESEGSWVDSFARIYVKVPPVLE---EGWE 2075
Qy 98 IIKELTSP-RLIKSHLPYRFLPSDLHNGDSKVIYMARPKDLVSVYQFHSRLR-TMSYR 155
Db 2076 MOKKGTQNFREIIVTMLREWAQDFESK-----LIRNFRLLRLQYSGIRBIRAMSQT 2131
Qy 156 GTFQFECRRFMNDKLG-----GSWFHVQE-----FWEHRMDSNVFLFKYEDMHR 201
Db 2132 VYFHERNEKVDTFVLVIQIRELTVQFHETEAILKRLGW--KLNNRIFFQHPDLMR 2189
Qy 202 -----DLVTWVEQLARFLGVSC--DKAQ 222
Db 2190 LLSVHENVMSIMMILTAQQGTVEHEGDELKEKAPIKDAEMVWVACSRFLCYFCRTSRQN 2249
Qy 223 LEALTEHCHQLVQCCNAEALPVGRGVLWKDIFTVSNNKEFDLVYKQK-----MG 274
Db 2250 QKAMFEHLFLSDNATMLLARSGLRSGVPLDVAYSSFMDNNELALKEELDQVAVYLS 2309
Qy 275 KCDDL 278
Db 2310 RCGL 2313

RESULT 63
T03851
thyroid hormone receptor activator molecule - human
C:Species: Homo sapiens (man)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004

C:Accession: T03851; T03749; T03443
R:Ikeshita, A.; Cardona, G.R.; Korbuchi, N.; Suen, C.S.; Chin, W.W.
J. Biol. Chem. 272, 27629-27634, 1997
A:Title: TRAM-1, a novel 160-kDa thyroid hormone receptor activator molecule, exhibits c
A:Reference number: Z15120; MUID:98010595; PMID:9346901
A:Accession: T03851
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-1424 <TAK>
A:Cross-references: UNIPROT:Q9Y6Q9; UNIPARC:UPI000012FB45; EMBL:AF016031; NID:G2584879;
R:Anick, S.L.; Kononen, J.; Walker, R.L.; Azores, D.O.; Tanner, M.M.; Guan, X.Y.; Sautc
Science 277, 965-968, 1997
A:Title: AIB1, a steroid receptor coactivator amplified in breast and ovarian Cancer.
A:Reference number: Z15053; MUID:97400625; PMID:9252329
A:Accession: T03749
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-1213, 1218-1424 <ANZ>
A:Cross-references: UNIPARC:UPI000002ABE3; EMBL:AF012108; NID:G2331249; PIDN:AAC51677.1;
R:Li, H.; Gomes, P.J.; Chen, J.D. 94, 8479-8484, 1997
Proc. Natl. Acad. Sci. U.S.A. 94, 8479-8484, 1997
A:Title: RAC3, a steroid/nuclear receptor-associated coactivator that is related to SRC
A:Reference number: Z14950; MUID:97385128; PMID:9238002
A:Accession: T03443
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-130, 'EA, 133-1213, 1218-1273, 1277-1424 <LIH>
A:Cross-references: UNIPARC:UPI0000169657; EMBL:AF010227; NID:G2318005; PIDN:AAC51663.1;
A:Experimental source: cell line HeLa
C:Genetics:
A:Gene: TRAM-1; AIB1; RAC3
A:Map position: 20

Query Match 6.6%; Score 100.5; DB 2; Length 1424;
Best Local Similarity 20.2%; Pred. No. 3.8;
Matches 52; Conservative 44; Mismatches 115; Indels 47; Gaps 11;
Qy 11 TPG-----EFESKYFPHGVRLPPFCRGKMEEIANFPVRPSDVMIVTPKSGT 58
Db 24 TPGQGLTCSGRKRREQESKYIE---ELABELISANLSDIDNFNVPDK-----C 69
Qy 59 SLLOEVVYLVSGQADPDEIGLWNIDEQLPVLEYQPQGLDII-KELTSRLIKSHLPYRFL 117
Db 70 AILKELTVRIQIKKEQK--TISNDDVDQKADVSTGGQVIDKDSGLGILLQALDGFILFV 127
Qy 118 PSDLHNGDSKVIYMARPKDLVSVYQF-----HRSRLTMSYRGTFQFCRRFMNDKLG 171
Db 128 V----NRDGNIVFVSEN---VTQYLOVKQEDLVNTSVYNILHEEDRKDFLKNLPKSTVN 179
Qy 172 YGSWFHVQEFWEHRMDSNVFLFKYEDMHRDLVTWVEQLARFLGVSC-DKAOLEALTEHC 230
Db 180 GVSMTNETQKSHTFNCRMLMKTTPHDILEDINASPENRQRYETMQCFALSQPRAMBEQ 239
Qy 231 HOLVDQC--CNAEALPVG 246
Db 240 EDL-QSCMICVARRITG 256

RESULT 64
T27665
hypochemical protein ZK1037.10 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
A:Accession: T27665
R:Baaham, V.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z20401
A:Accession: T27665
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-481 <WIL>
A:Cross-references: UNIPROT:O45992; UNIPARC:UPI0000061347; EMBL:Z81142; PIDN:CAB03509.1;
A:Experimental source: clone ZK1037

Db 236 V---VYLAGQVLKTDGSKIVDPRLQLNKKVREIKYSPGVTVKTEDNSVYSADYVMV 292

Qy 111 HLPVRLPSDLHGDSKVIYARNPKDLVSVYQPHRSRLRTMSYRGTFQEFRCRFMNDKL 170

Db 293 SASLGLVQSDL-----IQFKPLPTWKVRAIYQDMAV-----YTKILFKPRKFWPEGK 342

Qy 171 G-----YGSFHEVQEFWEHRMDSNVLFKYEDMHRDLVTWQLARFLGVS 217

Db 343 GREPFYASSRRGYGVW-----QEFKQYPDANVL-----LVTVTDEESRRIBQQ 388

Qy 218 CD---KAOL 223

Db 389 SDEQTKAEI 397

RESULT 68

T39468

anthranilate synthase (EC 4.1.3.27) component II - fission yeast (Schizosaccharomyces po

C:Species: Schizosaccharomyces pombe

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C:Accession: T39468

R:Mc Dougall, R.; Rajandream, M.A.; Barrell, B.G.; Seeger, K.; Quail, M.; Harris, D.

submitted to the EMBL Data Library, July 1999

A:Reference number: Z21856

A:Accession: T39468

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-759 <MCD>

A:Cross-references: UNIPROT:Q92370; UNIPARC:UPI0000013762D; EMBL:AL096874; PIDN:CAB51341.

A:Experimental source: strain 972h-; cosmid c1539

C:Genetics:

A:Gene: SPDB:SPBC1539.09c

A:Map position: 2

C:Superfamily: trpG-trpC-trpF trifunctional enzyme; trpC homology; trpF homology; trpG h

C:Keywords: carbon-carbon lyase; oxo-acid-lyase

Query Match 6.0%; Score 92; DB 2; Length 759;

Best Local Similarity 22.2%; Pred. No. 8.8;

Matches 42; Conservative 30; Mismatches 71; Indels 46; Gaps 7;

Qy 1 MAESEAEPTSP-GEFESKYFEFGVRL-----PPFCRGKWEETANFPVPSDV-W 49

Db 239 MRENSVSVSTKIRKQESILSKIAHQRLIDTAESKRKPGLSVGDLTQTVLNINAPPICNF 298

Qy 50 IVTPYKSGTSLQEVVYLVSGADPDEIGLNMIDEQ--LPVLEYPPQGLDIIKELTSRL 107

Db 299 YERLQSKPALMAEV-----KRASESK-GDIKDANAALQALTYAQVGASVISVITEPKW 352

Qy 108 IKSHL-----PYRFLPSDLHGDSKVIYARNPKDLVV 140

Db 353 FKGSINLDFARKAVEHVANRPAILRKDFIIDPYQIMEARLNGADSVLLIVAMLSREQLE 412

Qy 141 SYVQFHRSL 149

Db 413 SLYRFSKSL 421

RESULT 69

G71232

hypothetical protein PH0121 - Pyrococcus horikoshii

C:Species: Pyrococcus horikoshii

C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 21-Jul-2000

C:Accession: G71232

R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin

M.; Ohtuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi

DNA Res. 5, 55-76, 1998

A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a

A:Reference number: A71000; MUID:98344137; PMID:9679194

A:Accession: G71232

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-1434 <KAW>

A:Cross-references: UNIPARC:UPI0000015A654; GB:AP000001; NID:g3236128; PIDN:BAA29190.1; H

A:Experimental source: strain OT3

A:Note: this accession replaces an interim accession for a sequence replaced by GenBank

C:Genetics:

A:Gene: PH0121

Query Match 6.0%; Score 91.5; DB 2; Length 1434;

Best Local Similarity 22.6%; Pred. No. 23;

Matches 58; Conservative 46; Mismatches 86; Indels 67; Gaps 14;

Qy 45 PSD-VWIVTPYKSGTSLQEVVYLVSGADPDEIGLM-NIDEQLPVLEYPOP----GLDI 98

Db 1095 PEDFVFSLNACKYHTVIINE--NIVTHQCDDGDAVMLLLDALLNFSRYILPEKEGG--- 1149

Qy 99 IKELTSRLIKSHLPYRFLPSDLHGND-----SKVIYARNPKDLVSVYQFHRSL 149

Db 1150 --KMDAPLIVITTRLDPREVDSEVHMDIVRYYPLEFYEATYELKSPKELV----- 1197

Qy 150 RTMSYRGTFQFCFRFWMNDKLGYSWEFHVQEFWEHRMDSNVLFKLYE-DMHRDLVTMVE 208

Db 1198 -----GVIERVDRLKPEMYIYGLKFTH-----DTDDIALGPQMSLYKQLGDMEE 1242

Qy 209 QLARFLGVSCDKAQLEALTEHCHOLVDQCCNAEALPVGRGRVGLW--KDIFTVSNNEKF- 265

Db 1243 KVKQLDVA---RRIRAVDE--HKVAETILNHLIPDLRGNLRSFTRQEFRCVKCNTKER 1297

Qy 266 ----DLVYKQKMGKCDL 278

Db 1298 RPPLD-----GRCPI 1307

RESULT 70

S49879

hypothetical narbonin-like 2S protein - fava bean

C:Species: Vicia faba (fava bean)

C:Date: 05-Mar-1995 #sequence_revision 14-Jul-1995 #text_change 09-Jul-2004

C:Accession: S49879

R:Nong, V.; Schlesier, B.; Muentz, K.

submitted to the EMBL Data Library, November 1994

A:Description: The narbonin gene from Vicia faba L.

A:Reference number: S49849

A:Accession: S49879

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-285 <NON>

A:Cross-references: UNIPROT:Q41660; UNIPARC:UPI000000A571C; EMBL:Z46826; NID:g600101; PID

C:Superfamily: alcohol sulfotransferase

Query Match 5.9%; Score 91; DB 2; Length 285;

Best Local Similarity 25.1%; Pred. No. 2.8;

Matches 46; Conservative 25; Mismatches 80; Indels 32; Gaps 9;

Qy 13 GEFESKYF-EPHGVRLPPFCRGKMEET-----ANFPVRPSD--VWIVTPYKSGTS 59

Db 52 GNFEESMWDEFFGDPKVNKLTKHPEVKVVISIGRGVETFPDPAEQNIWVSNVAVKSLK 111

Qy 60 LLQEVVYLVSGADPDEIGL-----MNIDEQLPVLEYPOPGLDIIKELTSRLIKSHLPYR 115

Db 112 IIQK--YKNESGNLIDGIDINYEHIKSDAEPRLIG-----QLITELKKERDLNIHV-VS 163

Qy 116 FLPSDLHGDSKVIYARNPKDLVSVYQFHRSLRTMSYRGTFQEFRCRFMNDKLGYSW 175

Db 164 IAPSE--NNASSYLNLYNANPDINLVDYQFSNQLRHVSTEDAFVDIYKRVVND-----Y 216

Qy 176 FEH 178

Db 217 FTH 219

RESULT 71

S49898

hypothetical narbonin-like 2S protein (clone pVSNAG2) - spring vetch

C:Species: Vicia sativa (spring vetch, tare)

C:Date: 05-Mar-1995 #sequence_revision 14-Jul-1995 #text_change 09-Jul-2004

C;Accession: S49898
R;Nong, V.; Muentz, K.
submitted to the EMBL Data Library, November 1994
A;Description: A genomic sequence encoding putative narbonin from *Vicia sativa*.
A;Reference number: S49880
A;Accession: S49898
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-285 <N>
A;Cross-references: UNIPROT:Q41704; UNIPARC:UPI00000A97B2; EMBL:Z46835; NID:G600112; PID
C;Superfamily: alcohol sulfotransferase

Query Match 5.9%; Score 91; DB 2; Length 285;
Best Local Similarity 25.1%; Pred. No. 2.8;
Matches 46; Conservative 25; Mismatches 80; Indels 32; Gaps 9;

Qy 13 GEFESKYP-EFGVRLPFPFCRGKMEI-----ANFFVRPSD--VWIVTPKSGTS 59
Db 52 GNFEESWDEFFGPKVKNLTKHPEVKVVISIGRGVETFPDPAEQNIWVNAVKSLKL 111

Qy 60 LLQEVVYLVSGADPDEIGL-----WNIDEQPLVLEYPOFGLDIIKELTSPRLIKSHLPYR 115
Db 112 IIQK--YKNSGNLIDGIDINVEHIKSDAEPRLIG-----QLITELKKERDLNIHV-VS 163

Qy 116 FLPSDLHNGDSKVYMAENPKDLVSVYQFHRSLRTMSYRGTFQEFCEFRFNMNDKLYGWSW 175
Db 164 IAPSE-NNASSVLYNANPDINLDVYQFNSQLRHVSTEDAFVDIYKRVND-----Y 216

Qy 176 FEH 178
Db 217 FTH 219

RESULT 72
A;Accession: A71453
hypothetical protein PH0280 - *Pyrococcus horikoshii*
C;Species: *Pyrococcus horikoshii*
C;Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 05-Oct-2004
R;Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Hino, Y.; Yamamoto, S.; Sekin
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi
DNA Res. 5, 55-76, 1998
A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a
A;Reference number: A71000; MUID:98344137; PMID:9679194
A;Accession: A71453
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-508 <KAW>
A;Cross-references: UNIPROT:O58018; UNIPARC:UPI0000062D81; GB:AP0000001; NID:G3236128; PI
A;Experimental source: strain OT3
A;Note: this accession replaces an interim accession for a sequence replaced by GenBank
C;Genetics:
A;Gene: PH0280
C;Superfamily: *Methanococcus jannaschii* conserved hypothetical protein MJ0107

Query Match 5.9%; Score 91; DB 2; Length 508;
Best Local Similarity 25.5%; Pred. No. 6.2;
Matches 49; Conservative 26; Mismatches 65; Indels 52; Gaps 10;

Qy 40 NFPVRPSDVWIVTPK-SGTSLLQEVYLVSGADPDEIGLNN-----IDEQLPV 88
Db 150 DFPAR-ITAEIVDAPKLSLEGVLKANKYLLSEGADIVDIGVAGETNVKFLDVIPEFREV 208

Qy 89 L--EYP-----QFGLDIIEKLTLS-----PRLIKSHLPYRFLPSDLHNGDSKVYI 130
Db 209 LGREVPISLDSLNAKEAGLDFVDMILSDWSNVNVEDLVTDPKPVLIPTDMKNG-----Y 263

Qy 131 MARPKDLVSVYQFHRSLRTMSYRGTFQEFCEFRFNMNDKLYGWSWFEHVQEFWEHRMDSN 190
Db 264 PFLNPRDRVSELEKLEKALSIGY-----KRIADLI-----LEHYPGF-----SRSI 306

Qy 191 VLFLKYEDMHRD 202

Db 307 VAFSLYDRNRNE 318

RESULT 73
T46566
anthranilate synthase (EC 4.1.3.27), trifunctional - fission yeast (*Schizosaccharomyces*
N;Contains: glutamine amidotransferase (EC 2.6.1.-); indole-3-glycerol-phosphate synthase
C;Species: *Schizosaccharomyces pombe*
C;Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 09-Jul-2004
C;Accession: T46566
R;David, C.
submitted to the EMBL Data Library, October 1996
A;Reference number: Z23074
A;Accession: T46566
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-759 <DAV>
A;Cross-references: UNIPROT:Q92370; UNIPARC:UPI00001691BD; EMBL:Y09137; PIDN:CAA70348.1
A;Experimental source: strain 972(-)
C;Genetics:
A;Gene: trp-1
A;Map position: II
C;Superfamily: trpG-trpC-trpF trifunctional enzyme; trpC homology; trpF homology; trpG h
C;Keywords: aminotransferase; carbon-carbon lyase; carboxy-lyase; intramolecular oxidore

Query Match 5.9%; Score 91; DB 2; Length 759;
Best Local Similarity 22.2%; Pred. No. 11;
Matches 42; Conservative 29; Mismatches 72; Indels 46; Gaps 7;

Qy 1 MAESEAEPTSTP-GEFESKYPEFGVRL-----PPFCRGKMEETANFPVRPSDV-W 49
Db 239 MRENSVSSTKIRKQESILEKHAQRLIDTAESKRKPGLSVGLDTYVLNLIAPPICINF 298

Qy 50 IVTPYKSGTSLQEVYLVSGADPDEIGLNMIDEQ--LPVLEYPOFGLDIIEKLTSPRL 107
Db 299 YERLKQSKPALMAEV-----KRASPSK-GDIKLDANAIAQALTVAQVGASVISVLTPEKW 352

Qy 108 IKSHL-----PRFLPSDLHNGDSKVYMAENPKDLVV 140
Db 353 FKGSNLNDFVARKAVEHVANRPAILRKDFIIDPYEIMEARLNGADSVLLIVAMLSREQLE 412

Qy 141 SYVQFHRSL 149
Db 413 SLYFESKSL 421

RESULT 74
A57169
[heparan sulfate]-glucosamine N-sulfotransferase (EC 2.8.2.8) - human
N;Alternate names: N-heparan sulfate sulfotransferase (N-HSST)
N;Contains: desulfoheparin sulfotransferase (EC 2.8.2.8); glycosaminoglycan N-acetylgluc
C;Species: *Homo sapiens* (man)
C;Date: 23-Feb-1996 #sequence_revision 23-Feb-1996 #text_change 09-Jul-2004
C;Accession: A57169; G02129; G01581
R;Dixon, J.; Loftus, S.K.; Gladwin, A.J.; Scambler, P.J.; Wasmuth, J.J.; Dixon, M.J.
Genomics 26, 239-244, 1995
A;Title: Cloning of the human heparan sulfate-N-deacetylase/N-sulfotransferase gene from
A;Reference number: A57169; MUID:95324914; PMID:7601448
A;Accession: A57169
A;Molecule type: mRNA
A;Residues: 1-882 <DIX>
A;Cross-references: UNIPROT:P52848; UNIPARC:UPI000012CDA5; GB:U18918; NID:G976371; PIDN:
A;Note: authors translated the codon AAG for residue 42 as Leu, AAA for residue 110 as I
as Ile
R;Humphries, D.E.
submitted to the EMBL Data Library, September 1995
A;Reference number: G09202
A;Accession: G02129
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-882 <HOM>
A;Cross-references: UNIPARC:UPI000012CDA5; EMBL:U36600; NID:G1036796; PIDN:AAC27354.1; P
R;LaBell, T.L.

submitted to the EMBL Data Library, December 1994
A:Reference number: G07829
A:Accession: G01581
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-59, A', 61-688, 'G', 690-742, 'R', 744-882 <LAB>
A:Cross-references: UNIPARC:UPI000016A159; EMBL:U17970; NID:G841163; PIDN:AAA67765.1; PID:
C:Genetics:
A:Gene: GDB:HSST; NST1
A:Cross-references: GDB:593916; OMIM:600853
A:Map position: 5q32-5q33.1
C:Function: <DAG>
A:Description: as glycosaminoglycan N-acetylglucosaminyl N-deacetylase, hydrolyzes the N-
A:Pathway: heparan sulfate biosynthesis
C:Function: <NST>
A:Description: as desulfatase heparin sulfotransferase, catalyzes the formation of glucosamin
enosine-5'-phosphosulfate and converting it to adenosine 3',5'-bisphosphate
A:Pathway: heparan sulfate biosynthesis
C:Superfamily: Caenorhabditis elegans hypothetical protein F08B4.6
C:Keywords: glycoprotein; Golgi apparatus; hydrolase; sulfotransferase; transmembrane pr
F:1-17/Domain: intracellular #status predicted <INT>
F:18-39/Domain: transmembrane #status predicted <TRM>
F:40-882/Domain: trans-Golgi network luminal #status predicted <LUM>
F:231,351,401,667/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 5.9%; Score 90.5; DB 2; Length 882;
Best Local Similarity 20.9%; Pred. No. 15;
Matches 54; Conservative 36; Mismatches 95; Indels 73; Gaps 11;
Qy 12 PGEFSKVFHFHGVLPFPFCRGKMEIANFPVPSDVW-----IVTPKSGT 58
Db 563 PVQLAQKVFQIFSEKDPQWOPCED-----KRHKDWSKEKTCRDPFKLLIGPQKGT 617
Qy 59 SLLQEWVYLVSGADPD-----EIGLMN-----ID---EQLPVLE----- 90
Db 618 TAL-----VLF-LGMHPDLSNYPSSSETFEEIOFFNGHNYHKIDWMEFFPIPSNTTSD 672
Qy 91 YPQGLDIIKELTSPRLIKSHLPYRFLPSDLHGDSKVIYMARNPDKLVVSYQPHRSL- 149
Db 673 YFEKSANYFDSEVAPRRRAALLP-----KAKVTLILINPADRAYSWYQORAH 721
Qy 150 RTMSYRGTFQEB-----CRPFMNDKLGYSWFHFQEFWEHRMDSNVLFKLYEDM 199
Db 722 DPVALKYTFHEVITAGSDASSKLRALQNRCLVPGWYATHIERKWSAYHANQILVLDGKLL 781
Qy 200 HRDLVTMVEQLARFLGVS 217
Db 782 RTEPAKVMQVQKFLGVT 799
RESULT 75
D72379
hypothetical protein TM0428 - Thermotoga maritima (strain MSB8)
C:Species: Thermotoga maritima
C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C:Accession: D72379
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Swinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.
A>Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A:Reference number: A72200; MUID:99287316; PMID:10360571
A:Accession: D72379
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-435 <ARN>
A:Cross-references: UNIPROT:Q9WYQ9; UNIPARC:UPI0000003A0F; GB:AE001721; GB:AE000512; NID
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: TM0428
C:Superfamily: NADPH dehydrogenase chain OYE2

Query Match 5.9%; Score 90; DB 2; Length 435;

Best Local Similarity 21.0%; Pred. No. 6.1;
Matches 44; Conservative 42; Mismatches 64; Indels 60; Gaps 11;
Qy 58 TSLQEVVYLVSGADPDPEIGLMNIDEOLP-----VLEYPOP-----GL 96
Db 226 TKLLKDIIVLRKVEVIDITVRLNLSDFIPYPYGMGSTTEEGTLDTFEPCRLKKELEELGV 285
Qy 97 DIIKELTSPRLIKSHL--PY-----RFLPSDLHN--GDSKVIYMARNPDK-----LVVSY 142
Db 286 KIINVSTPYLKPINRPYDEMCKYNPPE-HPIVGAMRLNLAKEKLSKTLVWASG 344
Qy 143 YOFHSLRTMSYRGTFQEB-FCRRFMNDKLGYSWFHFQEFWEHRMDSNVLFKLYEDMHR 201
Db 345 FTWFRQFAPYVAGMLKNGWC-----DFVGFGR-----MTFAIPDPK 382
Qy 202 DLVTMVEQLARFLGVSCKD-KAQLEALTEHC 230
Db 393 DILTKGELDPKKVCITCNKCAELKAAGESC 412
RESULT 76
B89791
hypothetical protein SA0261 [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: B89791
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A>Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: B89791
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-332 <KUP>
A:Cross-references: UNIPROT:Q99WV4; UNIPARC:UPI00000054670; GB:BA000018; PID:gl3700187;
A:Experimental source: strain N315
C:Genetics:
A:Gene: SA0261
C:Superfamily: lac repressor

Query Match 5.8%; Score 89; DB 2; Length 332;
Best Local Similarity 25.4%; Pred. No. 5.2;
Matches 36; Conservative 16; Mismatches 38; Indels 52; Gaps 6;
Qy 24 GVRLP----PFCRGKMEBIANFPVRPSDVWITVTPKSGTSLQEVVYLVSGADPEIGL 79
Db 66 GVILPSLTNPFPSALMQSIHDH--KPSVDVICFLTSTATDLYDNIKHLIDRID----- 117
Qy 80 MNIDEQLPVLEYPOPGLDIIKELTSP-----RLIKSHLPYRFLPSDLHGDSKVIYMAR 134
Db 118 -----GLIIAQYISSPDALNNYVKKHVPYVVLVDQNDHOG----- 152
Qy 135 PKDLVSYQPHRSLRTMSYRG 156
Db 153 -----YTDF-----VRTNEYQG 164
RESULT 77
A42855
N-heparan sulfate sulfotransferase - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 20-Jun-2000
C:Accession: A42855
R:Hashimoto, Y.; Orellana, A.; Gil, G.; Hirschberg, C.B.
J. Biol. Chem. 267, 15744-15750, 1992
A>Title: Molecular cloning and expression of rat liver N-heparan sulfate sulfotransferas
A:Reference number: A42855; MUID:92348437; PMID:1379236
A:Accession: A42855
A:Status: preliminary
A:Molecule type: mRNA; protein
A:Residues: 1-882 <HAS>

Db 159 EIHQVLL-----QVADPPEKVASVIDHTYPN-----KQVVQLISSQIDADRMDDL 204

Qy 81 NIDEQLPVLETPQGLDIIKELTSRL:KSHLPYRFLPSDLHNGDSKVIYMARPKDLVV 140

Db 205 LRDSYFTGASYGE--FDLTRILVRPIENGIAFO-----RNGMHAI-----EDYVL 249

Qy 141 SYVQ-----FHRSLRTSYRGTFQFCRR-----FMNDKLGVSWFHVFQWFEHRMDSN 190

Db 250 SRYQMYMQVYHPATRAVEV--LLQNLLKRAKELYPEDKDFARTSPHLLFFFEKN----- 303

Qy 191 VLFLKYEDMRDLVTMVQLARFLGVSCKAQL-----EALTECHQOLV----- 234

Db 304 -----VLTLDYALDDCGVMNTYFQLWMTSPDKILADLSHRFVNKRKVPKSTTF 350

Qy 235 -----DQCNAEALPVGRVGLWKDIFTVSMNEKFDLVY 269

Db 351 SOEQDQDLASMRKLVED---IGFDDPYDT-AIHKNFIDLPI 386

RESULT 84

A49733

[heparan sulfate]-glucosamine N-sulfotransferase (EC 2.8.2.8) - mouse

N:Alternate names: N-heparan sulfate sulfotransferase (N-HSST)

N:Contains: desulfheparin sulfotransferase (EC 2.8.2.8); glucosaminoglycan N-acetylgluc

C:Species: Mus musculus (house mouse)

C>Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 09-Jul-2004

C:Accession: A49733; A53736

R:Orellana, A.; Hirschberg, C.B.; Wei, Z.; Swiedler, S.J.; Ishihara, M.

J. Biol. Chem. 269, 2270-2276, 1994

A:Title: Molecular cloning and expression of a glucosaminoglycan N-acetylglucosaminyl N-

A:Reference number: A49733; MUID:94124588; PMID:8294485

A:Accession: A49733

A>Status: Preliminary

A:Molecule type: mRNA

A:Residues: 1-883 <ER>

A:CROSS-references: UNIPROT:PS2850; UNIPARC:UPI0000000997; GB:U02304; NID:G457943; PIDN:

R:Eriksson, I.; Sandbaeck, D.; Ek, B.; Lindahl, U.; Kjellen, L.

J. Biol. Chem. 269, 10438-10443, 1994

A:Title: cDNA cloning and sequencing of mouse mastocytoma glucosaminyl N-deacetylase/N-

A:Reference number: A53736; MUID:94193735; PMID:8144627

A:Accession: A53736

A>Status: Preliminary

A:Molecule type: mRNA

A:Residues: 1-117, 'V', 119-395, 397-883 <ERI>

A:CROSS-references: UNIPARC:UPI0000163857; GB:X75885; NID:G474430; PIDN:CAA53479.1; PID:

C:Superfamily: Caenorhabditis elegans hypothetical protein F08B4.6

C:Keywords: carboxylic ester hydrolase; glycoprotein; sulfotransferase; transmembrane pr

Query Match 5.7%; Score 87.5; DB 2; Length 883;

Best Local Similarity 20.2%; Pred. No. 27; Mismatches 109; Indels 53; Gaps 9;

Matches 50; Conservative 36;

Qy 12 PQEFESKVFHFHGVRLPFCRGKMBEIANFPVRPSDVW-----IVTYPKSGT 58

Db 562 VPVLAQKYFELFPQSPSLMNPQCD-----KRHKDWSKENTCDRLPKFLVIGPKTGT 616

Qy 59 SLLQEVVL-----VSCGADP-----DIGLMN-----IDEQLPVLEYP-QPGLDIIKELT 103

Db 617 TAIHFPLSLHPAVTSFSPSPSTFEBIQFNGPNYHKIDWYMDFFVPFNSASTDFLFEKS 676

Qy 104 SPRLIKSHLPYR---FLPSDLHNGDSKVIYMARPKDLVVSYYQFHR-----L 149

Db 677 ATYFDSEVVPARGAALLPR-----AKITVLINPADRAYSWYQHRAGDPDIALNYTFY 730

Qy 150 RTMSYRGTFQFCRRFMNDKLGVSWFHVFQWFEHRMDSNVFLFKYEDMRDLVTMVQE 209

Db 731 QVISASSQAPLLRLSLQNRCLVPGVGYSTHLQRLWLYTPSGQLLIMDGQELRWPAASMEI 790

Qy 210 LARFLGVS 217

Db 791 IQKFLGIT 798

RESULT 85

C84500

probable retroelement pol polyprotein [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C:Accession: C84500

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Rounsley, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Unayam, L.; Tallon, L.

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: C84500

A>Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-976 <STO>

A:CROSS-references: UNIPARC:UPI000017A748; GB:AB002093; NID:G4557068; PIDN:AAD22507.1; G:

C:Genetics:

A:Gene: At2g12020

A:Map position: 2

Query Match 5.7%; Score 87.5; DB 2; Length 976;

Best Local Similarity 25.0%; Pred. No. 30;

Matches 47; Conservative 36; Mismatches 80; Indels 25; Gaps 12;

Qy 83 DEQLPVLEYPQGLDII-KELTSRLI---KSHLPYRFL-----PSDLHNGDSKVIYMARN 134

Db 8 DESKSSVEH-QRRNLILQKRLRLKLLSYWKQVLSIRPQTAVGPDVHRLOKAECQYQER 66

Qy 135 PKDLVVSYYQFHRSLRTWS--YRGTFQFCRRFMNDKLGVSWFHVFQWFE--HRMDS 189

Db 67 PFPLTI--YRSHVGERGMSIFTDIEDIMEVFMDSVYGSLEFDCLENLYKVLARCEE 124

Qy 190 NVFLKYEDMH---RDLVTWVEQLARFLGVSCKAQLAEALTECHQOLVDQCCNAEALPVG 246

Db 125 KHLVLNNEKCHFRVQDGVILGHRISEY-GIBADRAKIEVMTS--LQALD---NLKAVRSF 178

Qy 247 RGRVGLWK 254

Db 179 LGHAGFYR 186

RESULT 86

E75199

DNA-directed DNA polymerase (EC 2.7.7.7) II large chain PAB2404 - Pyrococcus abyssi (str

C:Species: Pyrococcus abyssi

C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 05-Oct-2004

C:Accession: E75199

R:anonymous, Genoscope

submitted to the EMBL Data Library, July 1999

A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru

A:Reference number: A75001

A>Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-1455 <KAW>

A:CROSS-references: UNIPROT:Q9V2F4; UNIPARC:UPI0000003472A; GB:AJ248283; GB:AL096836; NID:

A:Experimental source: strain Orsay

C:Genetics:

A:Gene: PAB2404

C:Superfamily: DNA polymerase II large subunit

C:Keywords: nucleotidyltransferase

Query Match 5.7%; Score 87.5; DB 2; Length 1455;

Best Local Similarity 22.9%; Pred. No. 52;

Matches 57; Conservative 46; Mismatches 89; Indels 57; Gaps 13;

Qy 37 EIANFP-VRPSD--VWIVTYPKSGTSLIQEVVVLVSQADPDEIGLM-NIDEQLPVLEYP 92

Db 1104 EIKIKYVRANDDFVSLNAKKYHNVIINE--NIVTHOCDCGDEDVAMLLDALLNFSRY 1161

Qy 93 QP-----GLDIIKELTSRLIKSHLPYRFLPSDLHNGD-----SKVIYMARPKDLV 139

Db 1162 LPEKRG-----KMDAPLVITTRLDPREVDSEVHNMIDIVRYYPLEFYEATYELKSPKELV 1216

Qy 140 VSYIQPHRSRLTMSYRGTFQEFRCRFEMNDKLGYSWFHEHVQBFWEHRMDSNVLFKYE-D 198

Db 1217 -----GVIERVEDRLGKPEMYGLKFTH-----DTDDIALGPKMS 1251

Qy 199 MHRDLVTWVEQLARFLGVSCDKRAQLAEALTEHCHQLVDQCCNAEALPVGRGRVGLW--KDI 256

Db 1252 LYKQLGDMEEKVRROLEVA---KRIRAVDE--HGVAEKILNGLHLPDLRGNLRSTFROBF 1306

Qy 257 FTVSNMEKF 265

Db 1307 RCVKCNTKF 1315

RESULT 87

A95150

C: conserved hypothetical protein SP1290 [imported] - Streptococcus pneumoniae (strain TIGR

C:Species: Streptococcus pneumoniae

C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004

C:Accession: A95150

R:Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid

on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,

nson, T.; Hickey, E.K.; Holt, I.E.

Science 293, 498-506, 2001

A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,

A:Title: Complete Genome Sequence of a Virulent Isolate of Streptococcus pneumoniae.

A:Reference number: A95000; MUID:21357209; PMID:11463916

A:Accession: A95150

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-434 <KUR>

A:Cross-references: UNIPROT:0970C9; UNIPARC:UPI00000517B0; GB:AE005672; PIDN:AAK75394.1;

A:Experimental source: strain TIGR4

C:Genetics:

A:Gene: SP1290

C:Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ1154

Query Match 5.7%; Score 86.5; DB 2; Length 434;

Best Local Similarity 21.8%; Pred. No. 12;

Matches 61; Conservative 38; Mismatches 98; Indels 83; Gaps 13;

Qy 21 EFHGVLRLPPFCRGKMEETIANFVRPSDVWIVTPKSGTSLLEQVVLVSQGADPEIGLM 80

Db 127 EIHQVLL-----QVAPDFPEKVASVIDHTYPN-----KQVQLISSQIDADRMDYL 172

Qy 81 NIDEQLPVLEYPPQGLDIKELTSRLIKSLHPYRFLPSDLHNGDSKVIYMARNPCKDLVV 140

Db 173 LRDSYFTGASYGE--FDLTRLIRVIRPIENGIAFQ-----RNGMHAI-----EDYVL 217

Qy 141 SYQQ-----FHSRLTMSYRGTFQEFRCR-----FMNDKLGYSWFHEHVQEFWEHRMDSN 190

Db 218 SRQYMYQVYFHPATRAFEV--LLQNLLKRAKELYPEDKDFARTSPHLLPFEKN---- 271

Qy 191 VLFLKYEDMHRDLVTWVEQLARFLGVSCKAQL-----EALTEHCHQLV----- 234

Db 272 -----VTLDYDALDDGVMTYFQLWMTSPDKILADLSHRFVNRKVKFSITF 318

Qy 235 -----DOCCNAEALPVGRGRVGLWKDIFTVSNMEKFDLYV 269

Db 319 SOEDQDQLTSMRKLVED---IGFDPDYIT-AIHKNFDLPY 354

RESULT 88

JDVLD

DNA-directed DNA polymerase (EC 2.7.7.7) - duck hepatitis virus

C:Species: duck hepatitis virus, DHBV

C:Date: 20-Sep-1984 #sequence_revision 08-Nov-1996 #text_change 09-Jul-2004

C:Accession: S12844; S36614; A00710

R:Mattes, F.; Tong, S.; Teubner, K.; Blum, H.E.

Nucleic Acids Res. 18, 6140, 1990

A:Title: Complete nucleotide sequence of a German duck hepatitis B virus.

A:Reference number: S12844; MUID:91045092; PMID:2235507

A:Accession: S12844

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-836 <MAT>

A:Cross-references: UNIPROT:066400; UNIPARC:UPI00001726E5; EMBL:X12798; NID:G59057

A:Experimental source: isolate DHBV FI-6

A:Note: this ORF is not annotated in GenBank entry DHBVFI6, release 103

R:Munshi, A.; Panda, S.K.

A:Description: Cloning sequencing and sequence comparison of the indian isolate.

A:Reference number: S36614

A:Accession: S36614

A:Molecule type: DNA

A:Residues: 1-35,'V',37-81,'I',83-115,'Y',117-176,'Y',178-184,'T',186-188,'Y',190-220,'T'

A:Cross-references: UNIPARC:UPI00000F76D8; EMBL:X74623; NID:G9397342; PIDN:CAAS2700.1; PI

A:Experimental source: isolate IDHBV

R:Mandart, E.; Kay, A.; Galibert, F.

J. Virol. 49, 782-792, 1984

A:Title: Nucleotide sequence of a cloned duck hepatitis B virus genome: comparison with

A:Reference number: A92997; MUID:84138772; PMID:6699938

A:Accession: A00710

A:Molecule type: DNA

A:Residues: 390-734,'A',736-836 <MAN>

A:Cross-references: UNIPARC:UPI00001726E6; GB:K01834

A:Note: only part of the sequence reported in the GenBank entry is shown in the publishe

C:Superfamily: hepatitis virus DNA-directed DNA polymerase

C:Keywords: DNA biosynthesis; nucleotidyltransferase

Query Match 5.6%; Score 86; DB 1; Length 836;

Best Local Similarity 19.3%; Pred. No. 33;

Matches 68; Conservative 55; Mismatches 119; Indels 110; Gaps 18;

Qy 3 ESEATPSTPGTEFESKYFEFHGVLRLPPFCRGKMEETIANFVRPSDVWIVTPKSGTSLIQ 62

Db 383 EAATRGRTSGKSVSPR---DSSAIPVRTSGASDK--NSPLEEENVVLRTGNTSPWNRIT 437

Qy 63 EVVYLVSQADPEIGLWIDBQLPVLEYQ-----PGLDIIEKLT---S 104

Db 438 GKFLVLDKNS-----RNTEARLVDFSQFSKGNAMRFPYVWSPNLSLTRLRLPVGM 490

Qy 105 PRLIKS-----HLPY-----RFLPSDLHNGDSKVIYMARNPCKDLVSYVQFHRSLRTM 152

Db 491 PRISLDLSQAYHLPLNPASSRLAVSD----GQRVYFRKAPMGVGLSPFLHLHFTTAL 546

Qy 153 SYRGTFQBFRRF-----MND-----KLGYGSWFHFVQ--- 180

Db 547 G-----SEISRRFNVTFTYMDDFLLCHPNARHLNAISHAVCSFLQELGIRINFDKTTPS 601

Qy 181 -----EFWEHRMDSNVLFKYEDMH-RDLVTWVEQLARFLGVSCKAQLAEALTEHCHQLV 234

Db 602 PVNEIRFLGYQIDEN--FMKIEESRWKELRTVIKKIK--VGEWYDWKCIQRFVGHNFV- 656

Qy 235 DOCCNAEALPVGRGRVGLWKDIF---TVSMNEKFD-----LVYKQKMGKCDL 278

Db 657 -----LPFTKGNIEMLKPMYAAITNQVNFSSFSSTRTLLYLKLTMGVCKL 700

RESULT 89

T29107

polymorphic antigen p150 precursor - Theileria parva

C:Species: Theileria parva

C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004

C:Accession: T29107

R:Skilton, R.; Iams, K.P.; Macklin, M.D.; Goblright, E.

A:Description: Characterisation of a polymorphic 150 kilodalton antigen of Theileria pa

A:Reference number: Z20571

A:Accession: T29107

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-1452 <SKI>

A:Cross-references: UNIPROT:Q27028; UNIPARC:UPI00000074921; EMBL:L47230; NID:G986990; PID

A:Experimental source: strain Muguga; ssp. parva


```

Qy 102 LTSPRLTKSHLPYRFL--PSDLNGDSKVIYMAKNPKDLVVSYQFHRSLRTWSYRGTFQ 159
      ::::: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 729 YEESVVAARISYPPVVVRPSYVLGRAMEI-----VYSDEELERYNYTYAVQIE 776
      ::::: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 160 E-----FCRRFPMD-----KLGVSWFPEHVQEFWEHRWDSNVLFLLKYEDMHR 201
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 777 PDHPILDKFLENAIEVDVDSLTDSTGKVVIGSIEMEIEAGIHGDS-ACSIPTYISLSD 835
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 202 DLVTMV-----EQLARFLGV 216
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 836 NVLTTRQWTEQLARALNV 854
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 96
D84514
dynam-in-like protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 31-Dec-2004
C:Accession: D84514
M.:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: D84514
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-782 <STO>
A:Cross-references: UNIPROT:Q9S147; UNIPARC:UPI0000178828; GB:AE002093; NID:g4587686; PT:
C:Genetics:
A:Gene: At3g14120
A:Map position: 2

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.8
Copyright (c) 1993 - 2006 Bioacceleration Ltd.
OM protein - protein search, using sw model
Run on: May 17, 2006, 11:05:06 ; Search time 51 Seconds
(without alignments)
487.425 Million cell updates/sec
Title: US-10-768-158-2
Perfect score: 1530
Sequence: 1 MAESEAEPTSPGPESEKYF.....FDLVYKQKMGKDLTFDFYL 284
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 650591 seqs, 87530628 residues
Total number of hits satisfying chosen parameters: 650591
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries
Database : Issued Patents AA: *
1: /EMC_Celerra_IDS33/ptodata/2/iaa/5_COMB.pep:*
2: /EMC_Celerra_IDS33/ptodata/2/iaa/6_COMB.pep:*
3: /EMC_Celerra_IDS33/ptodata/2/iaa/7_COMB.pep:*
4: /EMC_Celerra_IDS33/ptodata/2/iaa/H_COMB.pep:*
5: /EMC_Celerra_IDS33/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /EMC_Celerra_IDS33/ptodata/2/iaa/RE_COMB.pep:*
7: /EMC_Celerra_IDS33/ptodata/2/iaa/backfiles.pep:*
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	1530	100.0	284	US-09-786-240-7 Sequence 7, Appli
2	472	30.8	295	US-09-854-122-21 Sequence 21, Appli
3	456	29.8	295	US-09-328-174A-3 Sequence 3, Appli
4	449	29.3	296	US-09-949-016-6018 Sequence 6018, Ap
5	449	29.3	296	US-09-792-695A-2 Sequence 2, Appli
6	435	28.4	309	US-09-949-016-10328 Sequence 10328, A
7	434	28.4	294	US-08-325-562-2 Sequence 2, Appli
8	434	28.4	294	US-08-437-795-2 Sequence 2, Appli
9	430	28.1	295	US-09-150-133-13 Sequence 13, Appli
10	430	28.1	295	US-09-150-141-13 Sequence 13, Appli
11	430	28.1	295	US-09-374-493-13 Sequence 13, Appli
12	430	28.1	295	US-09-374-824-13 Sequence 13, Appli
13	430	28.1	295	US-09-374-492-13 Sequence 13, Appli
14	430	28.1	295	US-09-785-343-13 Sequence 13, Appli
15	430	28.1	295	US-10-411-976-13 Sequence 13, Appli
16	430	28.1	295	US-10-201-525-13 Sequence 13, Appli
17	422	27.6	304	US-09-609-816-5 Sequence 5, Appli
18	422	27.6	304	US-10-199-334-5 Sequence 5, Appli
19	422	27.6	304	US-10-199-329-5 Sequence 5, Appli
20	422	27.6	304	US-10-199-330-5 Sequence 5, Appli
21	419.5	27.4	283	US-09-609-816-8 Sequence 8, Appli
22	419.5	27.4	283	US-09-609-816-9 Sequence 9, Appli
23	419.5	27.4	283	US-10-199-334-8 Sequence 8, Appli
24	419.5	27.4	283	US-10-199-334-9 Sequence 9, Appli
25	419.5	27.4	283	US-10-199-329-8 Sequence 8, Appli
26	419.5	27.4	283	US-10-199-329-9 Sequence 9, Appli

27	419.5	27.4	283	US-10-199-330-8 Sequence 8, Appli
28	419.5	27.4	283	US-10-199-330-9 Sequence 9, Appli
29	408.5	26.7	283	US-09-609-816-10 Sequence 10, Appli
30	408.5	26.7	283	US-10-199-334-10 Sequence 10, Appli
31	408.5	26.7	283	US-10-199-329-10 Sequence 10, Appli
32	408.5	26.7	283	US-10-199-330-10 Sequence 10, Appli
33	407	26.6	304	US-09-609-816-6 Sequence 6, Appli
34	407	26.6	304	US-10-199-334-6 Sequence 6, Appli
35	407	26.6	304	US-10-199-329-6 Sequence 6, Appli
36	407	26.6	304	US-10-199-330-6 Sequence 6, Appli
37	401	26.2	304	US-09-609-816-4 Sequence 4, Appli
38	401	26.2	304	US-10-199-334-4 Sequence 4, Appli
39	401	26.2	304	US-10-199-329-4 Sequence 4, Appli
40	401	26.2	304	US-10-199-330-4 Sequence 4, Appli
41	381	24.9	350	US-09-949-016-6526 Sequence 6526, Ap
42	381	24.9	377	US-09-949-016-11211 Sequence 11211, A
43	380	24.8	264	US-09-949-016-8464 Sequence 8464, Ap
44	333	21.8	350	US-08-852-481-2 Sequence 2, Appli
45	303.5	19.8	303	US-09-795-926-2 Sequence 2, Appli
46	303.5	19.8	303	US-10-364-774-2 Sequence 2, Appli
47	292.5	19.1	265	US-09-795-926-6 Sequence 6, Appli
48	292.5	19.1	265	US-10-364-774-6 Sequence 6, Appli
49	275.5	18.0	320	US-09-854-122-20 Sequence 20, Appli
50	265.5	17.4	324	US-09-854-122-18 Sequence 18, Appli
51	259.5	17.0	302	US-09-854-122-19 Sequence 19, Appli
52	254	16.6	214	US-09-795-926-16 Sequence 16, Appli
53	254	16.6	214	US-10-364-774-16 Sequence 16, Appli
54	187	12.2	186	US-09-795-926-12 Sequence 12, Appli
55	187	12.2	186	US-10-364-774-12 Sequence 12, Appli
56	187	12.2	331	US-09-854-122-17 Sequence 17, Appli
57	187	12.2	391	US-09-854-122-16 Sequence 16, Appli
58	176	11.5	148	US-09-795-926-8 Sequence 8, Appli
59	176	11.5	148	US-10-364-774-8 Sequence 8, Appli
60	150.5	9.8	106	US-10-094-749-2841 Sequence 2841, Ap
61	137.5	9.0	97	US-09-795-926-18 Sequence 18, Appli
62	137.5	9.0	97	US-10-364-774-18 Sequence 18, Appli
63	115	7.5	104	US-09-513-999C-8076 Sequence 8076, Ap
64	111	7.3	42	US-09-167-681-49 Sequence 49, Appli
65	106.5	7.0	325	US-10-126-279-25 Sequence 25, Appli
66	106.5	7.0	325	US-10-286-606-25 Sequence 25, Appli
67	106.5	7.0	326	US-10-891-383A-25 Sequence 25, Appli
68	104.5	6.8	148	US-09-795-926-10 Sequence 10, Appli
69	104.5	6.8	148	US-10-364-774-10 Sequence 10, Appli
70	104.5	6.8	1402	US-09-125-635-12 Sequence 12, Appli
71	101	6.6	42	US-09-167-681-41 Sequence 41, Appli
72	100.5	6.6	1420	US-09-125-635-4 Sequence 4, Appli
73	99	6.5	32	US-09-167-681-34 Sequence 34, Appli
74	99	6.5	32	US-09-167-681-42 Sequence 42, Appli
75	99	6.5	32	US-09-167-681-50 Sequence 50, Appli
76	97.5	6.4	907	US-09-949-016-10609 Sequence 10609, A
77	94	6.1	36	US-09-167-681-36 Sequence 36, Appli
78	94	6.1	37	US-09-167-681-44 Sequence 44, Appli
79	94	6.1	37	US-09-167-681-52 Sequence 52, Appli
80	94	6.1	42	US-09-167-681-33 Sequence 33, Appli
81	93.5	6.1	110	US-09-795-926-4 Sequence 4, Appli
82	93.5	6.1	110	US-10-364-774-4 Sequence 4, Appli
83	93.5	6.1	254	US-09-586-106D-69 Sequence 69, Appli
84	93.5	6.1	254	US-10-799-870-69 Sequence 69, Appli
85	92.5	6.0	883	US-09-949-016-6398 Sequence 6398, Ap
86	92.5	6.0	892	US-09-949-016-7255 Sequence 7255, Ap
87	88	5.8	121	US-09-270-767-41537 Sequence 41537, A
88	88	5.8	610	US-09-949-016-8271 Sequence 8271, Ap
89	88	5.8	638	US-09-949-016-8270 Sequence 8270, Ap
90	88	5.8	702	US-09-949-016-7288 Sequence 7288, Ap
91	88	5.8	752	US-08-244-189-2 Sequence 2, Appli
92	87.5	5.7	172	US-09-345-473E-12 Sequence 12, Appli
93	87.5	5.7	172	US-09-862-027-12 Sequence 12, Appli
94	87	5.7	154	US-08-946-329A-49 Sequence 49, Appli
95	87	5.7	154	US-09-562-914-49 Sequence 49, Appli
96	86.5	5.7	499	US-09-248-796A-14484 Sequence 14484, A
97	86	5.6	1402	US-03-445-33E-2 Sequence 2, Appli
98	85.5	5.6	241	US-10-094-749-2510 Sequence 2510, Ap
99	85.5	5.6	287	US-10-104-047-2485 Sequence 2485, Ap

```

100      84      5.5      254      2      US-09-586-106D-77      Sequence 77, Appl
ALIGNMENTS
RESULT 1
US-09-786-240-7
; Sequence 7, Application US/09786240
; Patent No. 6558935
; GENERAL INFORMATION:
; APPLICANT: INCYTE PHARMACEUTICALS, INC.
; APPLICANT: TANG, Y. Tom
; APPLICANT: CORLEY, Neil C.
; APPLICANT: GUEGLER, Karl J.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: LAL, Preeti
; APPLICANT: YUE, Henry
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: AZIMZAI, Yalda
; TITLE OF INVENTION: HUMAN TRANSFERASE PROTEINS
; FILE REFERENCE: PF-0592 PCT
; CURRENT APPLICATION NUMBER: US/09/786.240
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/150,657; unassigned; 09/186,779; unassigned; 60/133,642
; PRIOR FILING DATE: 1998-09-10; 1998-09-10; 1998-09-10; 1998-11-04; 1999-05-11
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PERL Program
; SEQ ID NO 7
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6558935 1784742CD1
US-09-786-240-7

```

Query Match	100.0%	Score 1530;	DB 2;	Length 284;
Best Local Similarity	100.0%;	Pred. No. 3.9e-166;		
Matches 284;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy 1	MASEAETPTGFEFESKYFEFHGVRLP	PPFCRGKMEERIANPVPSDVWIVTPKSGTSL	60	
Db 1	MAESEAETPTGFEFESKYFEFHGVRLP	PPFCRGKMEERIANPVPSDVWIVTPKSGTSL	60	
Qy 61	LQEVVYLVUSQADPDEIGLMNIDQLPVL	EYVQPQGLDIIKELTSPRLIKSHLPFRFLPSD	120	
Db 61	LQEVVYLVUSQADPDEIGLMNIDQLPVL	EYVQPQGLDIIKELTSPRLIKSHLPFRFLPSD	120	
Qy 121	LHNGDSKVIYMARNPKDLVVSYYQFHRS	LRMTSYRGTFQEFRCRFMNDKLGYSWFHVQ	180	
Db 121	LHNGDSKVIYMARNPKDLVVSYYQFHRS	LRMTSYRGTFQEFRCRFMNDKLGYSWFHVQ	180	
Qy 181	EFWEHRMDSNVFLFKYEDMHRDLVTMVEQ	LARFLGVSCDKAQLEALTEHCHQLVDQCCNA	240	
Db 181	EFWEHRMDSNVFLFKYEDMHRDLVTMVEQ	LARFLGVSCDKAQLEALTEHCHQLVDQCCNA	240	
Qy 241	EALPVGRGRVGLWKDIFTVSNNKFDLVIYKQ	KMGKCDLTFDFPYL	284	
Db 241	EALPVGRGRVGLWKDIFTVSNNKFDLVIYKQ	KMGKCDLTFDFPYL	284	

RESULT 2
US-09-854-122-21
/ Sequence 21, Application US/09854122
/ Patent No. 6841718
/ GENERAL INFORMATION:
/ APPLICANT: ALBERTE, ROBERT
/ APPLICANT: SMITH, RANDALL S.
/ TITLE OF INVENTION: TRANSGENIC PLANTS INCORPORATING TRAIT OF ZOSTERA MARINA
/ FILE REFERENCE: PHA-007.01
/ CURRENT APPLICATION NUMBER: US/09/854.122
/ CURRENT FILING DATE: 2001-09-10

```

; PRIOR APPLICATION NUMBER: 60/202,529
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 295
; TYPE: PR1
; ORGANISM: Homo sapiens
; US-09-854-122-21

Query Match      30.8%; Score 472; DB 2; Length 295;
Best Local Similarity 35.8%; Pred.No.3.5e-45;
Matches 100; Conservative 52; Mismatches 101; Indels 26; Gaps 4

Qy 24 GVRLLPPRCRKMRKREIANPPRPSPDVWIVTPKSGTSLLOEVVYLVSGQADPDEIGLNMID 83
Db 17 GVPLIKYFABALGQFQRFQRPDLNTTPKSGTTWVSQILDIMYQGGDLEKCNRAPIY 76

Qy 84 EQPLVLEYPQP---GLDITKELTSPRLIKSHLPYRFLPSPLDHNGDSKVIYMAENPKDLV 139
Db 77 VRVPFLEVNDGPEPSGLETLKDTPPRLIKSHLPALLPQTLLDQKVKVYVAENPKDVA 136

Qy 140 VSYIYQHRSRLTWSYRGTFQEFRCRPMNDKLGYSWEFHEQEFWEHRMDSNVLFKYEDEM 199
Db 137 VSYIYHFRMEKAHPPEQGTWDSFLKFMAGEVSYGSWYQHVQEMWELSRTHPEVLVLFYEDM 196

Qy 200 HRDLVTMTVEQLARFLGVSCDKAOLEALTEHC-----HQLVDQCCNAEA 242
Db 197 KENPKREIQKILFEVGRSLSEETWDMFVQHTSPKEMKKNPMNTVITVPQLMDHSIS---- 253

Qy 243 LP-VGRGRVGLWKDITFTVSNNKXPDLYVKRMGKCDLTF 280
Db 254 -PFWRKMGADGKTTFTTVAQNERFDADYAEKMGAGCSLSF 291

```

```

RESULT 3
US-09-328-174A-3
; Sequence 3, Application US/09328174A
; Patent No. 6448003
; GENERAL INFORMATION:
; APPLICANT: Guida, Marco
; APPLICANT: Kurth, Janice
; TITLE OF INVENTION: Genotyping Human Phenol Sulfotransferase
; TITLE OF INVENTION: (SIP2)
; FILE REFERENCE: 4389-6 (formerly SEQ-16P)
; CURRENT APPLICATION NUMBER: US/09/328,174A
; CURRENT FILING DATE: 1999-06-08
; PRIOR APPLICATION NUMBER: 09/328,174
; PRIOR FILING DATE: 1999-06-08
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 295
; TYPE: PRT
; ORGANISM: H. sapiens
US-09-328-174A-3

```

	Query Match	29.8%; Score 456; DB 2; Length 295;
	Best Local Similarity	33.7%; Pred.No. 2.3e-43;
	Matches	94; Conservative 58; Mismatches 101; Indels 26; Gaps 4;
Qy	24	GVLPLPFCRCGRMEIANFPVPSGVWLVTPKSGTSLQLQEVYLVLSQGADPEIGLWNID 83 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : Db 17 GVPLIKYFABALGFLQSFOARPPDLLISTYTPKSGTTWSQTLDIMYOGGDLEKCHRAPIF 76 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : Qy 84 EQLPVLEYPOP---GLDIIKELTSPLLIKSHLPYRELPSPDLNHGDSKVIVYAANPKDLV 139 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : Db 77 MRVPFLFKYPGGIPSGMETUKNTPAPRLKLTHLLPALLPOTLLDQKVIVVAENAKDVA 136 Qy 140 VSYVQHRSIRUWTSYRGTFQECRRFMNDKLGYSWFHEHQBEWEHRMDSNVLFKYEDM 199 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : Db 137 VSYYHFVMAKVVPHPGCTWSEFLKFWAGSVSYGSWYCHOEWELSKTHLPVLVLYEDM 196 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :


```

192 SWEKGKSPVLFYFEDLKEVILKHLERKPESELVDRIIHTHSFOEMKNPS 251
193
232 ----QLVQDCNABALP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
194
252 TNYTTLPDEIMNQKLSPFMRKGITGDKWKNHFTVALNEKFDKHVEQOMKESTLKF 305
195
RESULT 7
US-08-325-562-2
; Sequence 2, Application US/08325562
; Patent No. 5714594
; GENERAL INFORMATION:
; APPLICANT: Weinshilboun, Richard M.
; APPLICANT: Aksoy, Ibrahim A.
; APPLICANT: Wood, Thomas C.
; TITLE OF INVENTION: CDNA CLONING AND EXPRESSION OF HUMAN
; TITLE OF INVENTION: LIVER ESTROGEN SULFOTRANSFERASE
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg & Woessner P.A.
; STREET: 3500 IDS Center
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/325.562
; FILING DATE: 18-OCT-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Muetting, Ann M.
; REGISTRATION NUMBER: 33,977
; REFERENCE/DOCKET NUMBER: 150.131US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-339-0331
; TELEFAX: 612-339-3061
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 294 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-325-562-2

Query Match 28.4%; Score 434; DB 1; Length 294;
Best Local Similarity 33.3%; Pred. No. 7.6e-41;
Matches 95; Conservative 56; Mismatches 116; Indels 18; Gaps 3;

Qy 14 EFESKYFEFHGVRLPPFCRGKMEIANFPVRPSDVMTVTPKSGTSLQEVVYLVSQAD 73
Db 6 DYEFKEFEVHGILMYKDFVKVNDVNEAFQAPDDLVATYTPKSGTTWSEIVYMYKEGD 65
Qy 74 PDEIGLMNIDQLPVLEYQPQ-----GLDIIKELTSPRLIKSHLPYRFLPSPDLHNGDSKVI 129
Db 66 VEKCKEDVFNRIFFLECKENLMNGVKQLDEMNSPRIVKTHLPPELLPASFWEKDCKII 125
Qy 130 YMAENPKDLVSYVQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVEHVEHMD 189
Db 126 YLCRNAKDVAVSFYFFFLMAGHPNPGSPFPEFVEKFMQGVPPYGSWKYKHSWKEGKSP 185
Qy 190 NVFLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLAEALTEHCH-----QLVDQ 236
Db 186 RVLFYFEDLKEDIRKEVILKHLERKPESELVDRIIHTHSFOEMKNPSNTYTTLPDE 245
Qy 237 CCNAEALP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 246 IMNQKLSPFMRKGITGDKWKNHFTVALNEKFDKHVEQOMKESTLKF 290

RESULT 9
US-09-150-133-13
; Sequence 13, Application US/09150133B
; Patent No. 6060295
```

```

192 SWEKGKSPVLFYFEDLKEVILKHLERKPESELVDRIIHTHSFOEMKNPS 251
193
232 ----QLVQDCNABALP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
194
252 TNYTTLPDEIMNQKLSPFMRKGITGDKWKNHFTVALNEKFDKHVEQOMKESTLKF 305
195
RESULT 8
US-08-437-795-2
; Sequence 2, Application US/08437795
; Patent No. 5744355
; GENERAL INFORMATION:
; APPLICANT: Weinshilboun, Richard M.
; APPLICANT: Aksoy, Ibrahim A.
; APPLICANT: Wood, Thomas C.
; TITLE OF INVENTION: CDNA CLONING AND EXPRESSION OF HUMAN
; TITLE OF INVENTION: LIVER ESTROGEN SULFOTRANSFERASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg & Woessner P.A.
; STREET: 3500 IDS Center
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/437,795
; FILING DATE: 09-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Muetting, Ann M.
; REGISTRATION NUMBER: 33,977
; REFERENCE/DOCKET NUMBER: 150.137US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-339-0331
; TELEFAX: 612-339-3061
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 294 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-437-795-2

Query Match 28.4%; Score 434; DB 1; Length 294;
Best Local Similarity 33.3%; Pred. No. 7.6e-41;
Matches 95; Conservative 56; Mismatches 116; Indels 18; Gaps 3;

Qy 14 EFESKYFEFHGVRLPPFCRGKMEIANFPVRPSDVMTVTPKSGTSLQEVVYLVSQAD 73
Db 6 DYEFKEFEVHGILMYKDFVKVNDVNEAFQAPDDLVATYTPKSGTTWSEIVYMYKEGD 65
Qy 74 PDEIGLMNIDQLPVLEYQPQ-----GLDIIKELTSPRLIKSHLPYRFLPSPDLHNGDSKVI 129
Db 66 VEKCKEDVFNRIFFLECKENLMNGVKQLDEMNSPRIVKTHLPPELLPASFWEKDCKII 125
Qy 130 YMAENPKDLVSYVQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVEHVEHMD 189
Db 126 YLCRNAKDVAVSFYFFFLMAGHPNPGSPFPEFVEKFMQGVPPYGSWKYKHSWKEGKSP 185
Qy 190 NVFLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLAEALTEHCH-----QLVDQ 236
Db 186 RVLFYFEDLKEDIRKEVILKHLERKPESELVDRIIHTHSFOEMKNPSNTYTTLPDE 245
Qy 237 CCNAEALP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 246 IMNQKLSPFMRKGITGDKWKNHFTVALNEKFDKHVEQOMKESTLKF 290

RESULT 9
US-09-150-133-13
; Sequence 13, Application US/09150133B
; Patent No. 6060295
```

```

; GENERAL INFORMATION:
; APPLICANT: The Board of Regents of the University of Oklahoma
; TITLE OF INVENTION: TYROSYLPROTEIN SULFOTRANSFERASES, NUCLEIC ACIDS ENCODING
; FILE REFERENCE: 5820.504
; CURRENT APPLICATION NUMBER: US/09/150,133B
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: Wordperfect 5.1 (saved in ASCII format)
; SEQ ID NO 13
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: estrogen sulfotransferase
US-09-150-133-13

Query Match      28.1%; Score 430; DB 2; Length 295;
Best Local Similarity 33.0%; Pred. No. 2.2e-40;
Matches 94; Conservative 57; Mismatches 116; Indels 18; Gaps 3;

QY 14 EFESKYFEFHGVRLPPFCRCGMKEIANFPVPSDVMIIVTPKSGTSLQLQEVVYLVSQGAD 73
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 7 EYEVFGFGRGLVMDKRTKYWEDVEMFLARPDDLVIATYPKSGTTWSEVVYMIYKEGD 66

QY 74 PDEIGLMNIDQLPVLEYPOP-----GLDIIKELTSPRLIKSHLPYRFLPSPDLHNGDSKVI 129
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 67 VEKCKEDAIFNRIPLYLECRNEDLINGIKOLKEKESPRIVKTHLPKVLPPASFWEKNCXMI 126

QY 130 YMARNPDKLVVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVOEFWEHRMDS 189
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 127 YLCRNAKDVAVSYYYFLMLMITSYNPNKPSFSEFEVKFMQGVPPYGSWYDHWKAWKEKSKNS 186

QY 190 NVFLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLEALTEHCH-----QLVDQ 236
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 187 RVLFMFYEDMKEDIRREVVKLIIEFLERKPSAELVDRIIQTTSFQEMKNPNSTNYTMPEE 246

QY 237 CCNAEALP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 247 MNQKVSPFMRKGIIGDKWNHFPPEALRERFDEHYKQMKDCTVKF 291

; GENERAL INFORMATION:
; APPLICANT: The Board of Regents of the University of Oklahoma
; TITLE OF INVENTION: TYROSYLPROTEIN SULFOTRANSFERASES, AND METHODS OF USE THEREOF
; FILE REFERENCE: 5820.495
; CURRENT APPLICATION NUMBER: US/09/150,141B
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: Wordperfect 5.1 (saved in ASCII format)
; SEQ ID NO 13
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: estrogen sulfotransferase
US-09-150-141-13

Query Match      28.1%; Score 430; DB 2; Length 295;
Best Local Similarity 33.0%; Pred. No. 2.2e-40;
Matches 94; Conservative 57; Mismatches 116; Indels 18; Gaps 3;

QY 14 EFESKYFEFHGVRLPPFCRCGMKEIANFPVPSDVMIIVTPKSGTSLQLQEVVYLVSQGAD 73
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 7 EYEVFGFGRGLVMDKRTKYWEDVEMFLARPDDLVIATYPKSGTTWSEVVYMIYKEGD 66

QY 74 PDEIGLMNIDQLPVLEYPOP-----GLDIIKELTSPRLIKSHLPYRFLPSPDLHNGDSKVI 129
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 67 VEKCKEDAIFNRIPLYLECRNEDLINGIKOLKEKESPRIVKTHLPKVLPPASFWEKNCXMI 126

QY 130 YMARNPDKLVVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVOEFWEHRMDS 189
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 127 YLCRNAKDVAVSYYYFLMLMITSYNPNKPSFSEFEVKFMQGVPPYGSWYDHWKAWKEKSKNS 186

QY 190 NVFLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLEALTEHCH-----QLVDQ 236
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 187 RVLFMFYEDMKEDIRREVVKLIIEFLERKPSAELVDRIIQTTSFQEMKNPNSTNYTMPEE 246

QY 237 CCNAEALP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 247 MNQKVSPFMRKGIIGDKWNHFPPEALRERFDEHYKQMKDCTVKF 291

; GENERAL INFORMATION:
; APPLICANT: The Board of Regents of the University of Oklahoma
; TITLE OF INVENTION: TYROSYLPROTEIN SULFOTRANSFERASES, AND METHODS OF USE THEREOF
; FILE REFERENCE: 5820.495
; CURRENT APPLICATION NUMBER: US/09/150,141B
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: Wordperfect 5.1 (saved in ASCII format)
; SEQ ID NO 13
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: estrogen sulfotransferase
US-09-150-141-13

Query Match      28.1%; Score 430; DB 2; Length 295;
Best Local Similarity 33.0%; Pred. No. 2.2e-40;
Matches 94; Conservative 57; Mismatches 116; Indels 18; Gaps 3;

QY 14 EFESKYFEFHGVRLPPFCRCGMKEIANFPVPSDVMIIVTPKSGTSLQLQEVVYLVSQGAD 73
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 7 EYEVFGFGRGLVMDKRTKYWEDVEMFLARPDDLVIATYPKSGTTWSEVVYMIYKEGD 66

QY 74 PDEIGLMNIDQLPVLEYPOP-----GLDIIKELTSPRLIKSHLPYRFLPSPDLHNGDSKVI 129
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

```

```

Db 67 VEKCKEDAIFNRIPLYLECRNEDLINGIKOLKEKESPRIVKTHLPKVLPPASFWEKNCXMI 126
QY 130 YMARNPDKLVVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVOEFWEHRMDS 189
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 127 YLCRNAKDVAVSYYYFLMLMITSYNPNKPSFSEFEVKFMQGVPPYGSWYDHWKAWKEKSKNS 186
QY 190 NVFLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLEALTEHCH-----QLVDQ 236
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 187 RVLFMFYEDMKEDIRREVVKLIIEFLERKPSAELVDRIIQTTSFQEMKNPNSTNYTMPEE 246
QY 237 CCNAEALP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 247 MNQKVSPFMRKGIIGDKWNHFPPEALRERFDEHYKQMKDCTVKF 291

RESULT 11
US-09-374-493-13
; Sequence 13, Application US/09374493
; Patent No. 6204016
; GENERAL INFORMATION:
; APPLICANT: The Board of Regents of the University of Oklahoma
; TITLE OF INVENTION: TYROSYLPROTEIN SULFOTRANSFERASES AND METHODS OF USE THEREOF
; FILE REFERENCE: 5820.546
; CURRENT APPLICATION NUMBER: US/09/374,493
; CURRENT FILING DATE: 1999-08-13
; EARLIER APPLICATION NUMBER: 09/150,133
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 60/072,994
; EARLIER FILING DATE: 1998-04-29
; EARLIER APPLICATION NUMBER: PCT/US99/16750
; EARLIER FILING DATE: 1999-07-23
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: Wordperfect 8.0 (saved in ASCII format)
; SEQ ID NO 13
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: estrogen sulfotransferase
US-09-374-493-13

Query Match      28.1%; Score 430; DB 2; Length 295;
Best Local Similarity 33.0%; Pred. No. 2.2e-40;
Matches 94; Conservative 57; Mismatches 116; Indels 18; Gaps 3;

QY 14 EFESKYFEFHGVRLPPFCRCGMKEIANFPVPSDVMIIVTPKSGTSLQLQEVVYLVSQGAD 73
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 7 EYEVFGFGRGLVMDKRTKYWEDVEMFLARPDDLVIATYPKSGTTWSEVVYMIYKEGD 66

QY 74 PDEIGLMNIDQLPVLEYPOP-----GLDIIKELTSPRLIKSHLPYRFLPSPDLHNGDSKVI 129
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 67 VEKCKEDAIFNRIPLYLECRNEDLINGIKOLKEKESPRIVKTHLPKVLPPASFWEKNCXMI 126

QY 130 YMARNPDKLVVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVOEFWEHRMDS 189
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 127 YLCRNAKDVAVSYYYFLMLMITSYNPNKPSFSEFEVKFMQGVPPYGSWYDHWKAWKEKSKNS 186
QY 190 NVFLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLEALTEHCH-----QLVDQ 236
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 187 RVLFMFYEDMKEDIRREVVKLIIEFLERKPSAELVDRIIQTTSFQEMKNPNSTNYTMPEE 246
QY 237 CCNAEALP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 247 MNQKVSPFMRKGIIGDKWNHFPPEALRERFDEHYKQMKDCTVKF 291

RESULT 12
US-09-374-824-13
; Sequence 13, Application US/09374824
; Patent No. 6207414
; GENERAL INFORMATION:
; APPLICANT: The Board of Regents of the University of Oklahoma
; TITLE OF INVENTION: TYROSYLPROTEIN SULFOTRANSFERASES AND METHODS OF USE THEREOF

```

```

; FILE REFERENCE: 5820.547
; CURRENT APPLICATION NUMBER: US/09/374,824
; EARLIER FILING DATE: 1999-08-13
; EARLIER APPLICATION NUMBER: 09/150,133
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 60/072,994
; EARLIER FILING DATE: 1998-01-29
; EARLIER APPLICATION NUMBER: PCT/US99/16750
; EARLIER FILING DATE: 1999-07-23
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: Wordperfect 8.0 (saved in ASCII format)
; SEQ ID NO 13
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: estrogen sulfotransferase
US-09-374-824-13

Query Match      28.1%; Score 430; DB 2; Length 295;
Best Local Similarity 33.0%; Pred. No. 2.2e-40;
Matches 94; Conservative 57; Mismatches 116; Indels 18; Gaps 3;

QY 14 EFESKYFEFHGVRLLPFCRCGMEEIANFPVRPSDVWIVTPKSGTSLQLQEVVLYVSQAD 73
DB 7 EYEVFGFGRGLVMDKRFKTYKWEDEMFLARPDDLVIATYPKSGTTWISVVYMIYKEGD 66

QY 74 PDEIGLMNIDBQLPVLEYQPQ-----GLDIIKELTSPRLIKSHLPYRFLPDLHNGDSKVI 129
DB 67 VEKCKEDAIENRIPYLECRNEDLINGIKQLKESPRIVKTHLPKVLPAFWEKNCMI 126

QY 130 YMARNPDKLVVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVEHQRMS 189
DB 127 YLCRNAKDAVSYYYFLMITSYDNPKNKSFSEFVEKFMQGVPGYSWDYHVKAWKSKNS 186

QY 190 NVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAELTEHCH-----QLVDQ 236
DB 187 RVLFMFYEDMKEDIRREVVKLIEFLERKPSAELVDRIIQTFSFQEMKNPNSTNTMPEE 246

QY 237 CCNAEALP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
DB 247 MNQKVSPFMRKGIIGDKNHFPEALRERFDEHYKQKMKDCTVKF 291

RESULT 13
US-09-374-492-13
; Sequence 13, Application US/09374492
; Patent No. 6207432
; GENERAL INFORMATION:
; APPLICANT: The Board of Regents of the University of Oklahoma
; TITLE OF INVENTION: TYROSYLPROTEIN SULFOTRANSFERASES AND METHODS OF USE THEREOF
; FILE REFERENCE: 5820.545
; CURRENT APPLICATION NUMBER: US/09/374,492
; EARLIER FILING DATE: 1999-08-13
; EARLIER APPLICATION NUMBER: 09/150,141
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 60/072,994
; EARLIER FILING DATE: 1998-01-29
; EARLIER APPLICATION NUMBER: PCT/US99/16750
; EARLIER FILING DATE: 1999-07-23
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: Wordperfect 8.0 (saved in ASCII format)
; SEQ ID NO 13
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: estrogen sulfotransferase
US-09-374-492-13

Query Match      28.1%; Score 430; DB 2; Length 295;
Best Local Similarity 33.0%; Pred. No. 2.2e-40;
Matches 94; Conservative 57; Mismatches 116; Indels 18; Gaps 3;

QY 14 EFESKYFEFHGVRLLPFCRCGMEEIANFPVRPSDVWIVTPKSGTSLQLQEVVLYVSQAD 73
DB 7 EYEVFGFGRGLVMDKRFKTYKWEDEMFLARPDDLVIATYPKSGTTWISVVYMIYKEGD 66

QY 74 PDEIGLMNIDBQLPVLEYQPQ-----GLDIIKELTSPRLIKSHLPYRFLPDLHNGDSKVI 129
DB 67 VEKCKEDAIENRIPYLECRNEDLINGIKQLKESPRIVKTHLPKVLPAFWEKNCMI 126

QY 130 YMARNPDKLVVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVEHQRMS 189
DB 127 YLCRNAKDAVSYYYFLMITSYDNPKNKSFSEFVEKFMQGVPGYSWDYHVKAWKSKNS 186

QY 190 NVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAELTEHCH-----QLVDQ 236
DB 187 RVLFMFYEDMKEDIRREVVKLIEFLERKPSAELVDRIIQTFSFQEMKNPNSTNTMPEE 246

QY 237 CCNAEALP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
DB 247 MNQKVSPFMRKGIIGDKNHFPEALRERFDEHYKQKMKDCTVKF 291

RESULT 13
US-09-374-492-13
; Sequence 13, Application US/09374492
; Patent No. 6207432
; GENERAL INFORMATION:
; APPLICANT: The Board of Regents of the University of Oklahoma
; TITLE OF INVENTION: TYROSYLPROTEIN SULFOTRANSFERASES AND METHODS OF USE THEREOF
; FILE REFERENCE: 5820.545
; CURRENT APPLICATION NUMBER: US/09/374,492
; EARLIER FILING DATE: 1999-08-13
; EARLIER APPLICATION NUMBER: 09/150,141
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 60/072,994
; EARLIER FILING DATE: 1998-01-29
; EARLIER APPLICATION NUMBER: PCT/US99/16750
; EARLIER FILING DATE: 1999-07-23
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: Wordperfect 8.0 (saved in ASCII format)
; SEQ ID NO 13
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: estrogen sulfotransferase
US-09-374-492-13

Query Match      28.1%; Score 430; DB 2; Length 295;
Best Local Similarity 33.0%; Pred. No. 2.2e-40;
Matches 94; Conservative 57; Mismatches 116; Indels 18; Gaps 3;

```

```

QY 14 EFESKYFEFHGVRLLPFCRCGMEEIANFPVRPSDVWIVTPKSGTSLQLQEVVLYVSQAD 73
DB 7 EYEVFGFGRGLVMDKRFKTYKWEDEMFLARPDDLVIATYPKSGTTWISVVYMIYKEGD 66

QY 74 PDEIGLMNIDBQLPVLEYQPQ-----GLDIIKELTSPRLIKSHLPYRFLPDLHNGDSKVI 129
DB 67 VEKCKEDAIENRIPYLECRNEDLINGIKQLKESPRIVKTHLPKVLPAFWEKNCMI 126

QY 130 YMARNPDKLVVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVEHQRMS 189
DB 127 YLCRNAKDAVSYYYFLMITSYDNPKNKSFSEFVEKFMQGVPGYSWDYHVKAWKSKNS 186

QY 190 NVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAELTEHCH-----QLVDQ 236
DB 187 RVLFMFYEDMKEDIRREVVKLIEFLERKPSAELVDRIIQTFSFQEMKNPNSTNTMPEE 246

QY 237 CCNAEALP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
DB 247 MNQKVSPFMRKGIIGDKNHFPEALRERFDEHYKQKMKDCTVKF 291

RESULT 14
US-09-785-343-13
; Sequence 13, Application US/09785343
; Patent No. 6605455
; GENERAL INFORMATION:
; APPLICANT: The Board of Regents of the University of Oklahoma
; TITLE OF INVENTION: TYROSYLPROTEIN SULFOTRANSFERASES AND METHODS OF USE THEREOF
; FILE REFERENCE: 5927.003
; CURRENT APPLICATION NUMBER: US/09/785,343
; CURRENT FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: 09/150,133
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/072,994
; PRIOR FILING DATE: 1998-01-29
; PRIOR APPLICATION NUMBER: PCT/US99/16750
; PRIOR FILING DATE: 1999-07-23
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: Wordperfect 8.0 (saved in ASCII format)
; SEQ ID NO 13
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: estrogen sulfotransferase
US-09-785-343-13

Query Match      28.1%; Score 430; DB 2; Length 295;
Best Local Similarity 33.0%; Pred. No. 2.2e-40;
Matches 94; Conservative 57; Mismatches 116; Indels 18; Gaps 3;

QY 14 EFESKYFEFHGVRLLPFCRCGMEEIANFPVRPSDVWIVTPKSGTSLQLQEVVLYVSQAD 73
DB 7 EYEVFGFGRGLVMDKRFKTYKWEDEMFLARPDDLVIATYPKSGTTWISVVYMIYKEGD 66

QY 74 PDEIGLMNIDBQLPVLEYQPQ-----GLDIIKELTSPRLIKSHLPYRFLPDLHNGDSKVI 129
DB 67 VEKCKEDAIENRIPYLECRNEDLINGIKQLKESPRIVKTHLPKVLPAFWEKNCMI 126

QY 130 YMARNPDKLVVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVEHQRMS 189
DB 127 YLCRNAKDAVSYYYFLMITSYDNPKNKSFSEFVEKFMQGVPGYSWDYHVKAWKSKNS 186

QY 190 NVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAELTEHCH-----QLVDQ 236
DB 187 RVLFMFYEDMKEDIRREVVKLIEFLERKPSAELVDRIIQTFSFQEMKNPNSTNTMPEE 246

QY 237 CCNAEALP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
DB 247 MNQKVSPFMRKGIIGDKNHFPEALRERFDEHYKQKMKDCTVKF 291

```

```
RESULT 15
US-10-411-976-13
; Sequence 13, Application US/10411976
; Patent No. 6713283
; GENERAL INFORMATION:
; APPLICANT: Moore, Kevin L.
; TITLE OF INVENTION: TYROSYLPROTEIN SULFOTRANSFERASES
; FILE REFERENCE: 5864.027
; CURRENT APPLICATION NUMBER: US/10/411,976
; CURRENT FILING DATE: 2003-04-11
; PRIOR APPLICATION NUMBER: 09/785,343
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: Patencin version 3.1
; SEQ ID NO 13
; LENGTH: 295
; TYPE: PRT
; ORGANISM: estrogen sulfotransferase
US-10-411-976-13

Query Match      28.1%; Score 430; DB 2; Length 295;
Best Local Similarity 33.0%; Pred. No. 2.2e-40;
Matches 94; Conservative 57; Mismatches 116; Indels 18; Gaps 3;

QY 14 EFESKYFEFHGVRLPFCRCGKMEETANPPVRSDWIVTYPKSGTSLLOEVVLYVLSQGAD 73
DB 7 EYEVFGFGRGVLMDKRFTKYWEDVEMFLARPDDLVIAITYPKSGTTWISVVYMIYKSGD 66
QY 74 PDEIGLWNIIDEOLPVLEYPOP---GLDIIKELTSPRLIKSHLPYRFLPSPDLHNGDSKVI 129
DB 67 VEKCKEDALFNRIPLYECRNEDLINGIKOLAKESPRIVKTHLPKVLPAFWKNCXKI 126
QY 130 YMARNPKOLVSVYQFHRSLRMTSYRGTFQEFRCRPMNDKLGYSWFEHVQEFWEHRMDS 189
DB 127 YLCRNAKDVAVSYYYFLMLITSYPNPKSFSEFVEKFMQGVPGYSWYDHWKAWWKSXNS 186
QY 190 NVLFUKYEDMHRDLVTWVEQLARFLGVSCDKAQLAALTEHCH-----QLVDQ 236
DB 187 RVLFNFIYEDMKDIRREVVKLIIEFLERKPSAELVDRIIQTHTSFQEMKNPNSTNYTMPEE 246
QY 237 CCNAEALP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
DB 247 MNQKVSPFMRKGIIGDMKNHFPPEALRERFDEHYKQMKDCTVKF 291

RESULT 16
US-10-201-525-13
; Sequence 13, Application US/10201525
; Patent No. 7029890
; GENERAL INFORMATION:
; APPLICANT: Board of Regents of the University of Oklahoma
; TITLE OF INVENTION: TYROSYLPROTEIN SULFOTRANSFERASES AND METHODS OF USE THEREOF
; FILE REFERENCE: 5827.005
; CURRENT APPLICATION NUMBER: US/10/201,525
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/785,343
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: PCT/US99/16750
; PRIOR FILING DATE: 1999-07-23
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: Patencin version 3.1
; SEQ ID NO 13
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-201-525-13

Query Match      28.1%; Score 430; DB 3; Length 295;
Best Local Similarity 33.0%; Pred. No. 2.2e-40;
Matches 94; Conservative 57; Mismatches 116; Indels 18; Gaps 3;

QY 14 EFESKYFEFHGVRLPFCRCGKMEETANPPVRSDWIVTYPKSGTSLLOEVVLYVLSQGAD 73
DB 7 EYEVFGFGRGVLMDKRFTKYWEDVEMFLARPDDLVIAITYPKSGTTWISVVYMIYKSGD 66
QY 74 PDEIGLWNIIDEOLPVLEYPOP---GLDIIKELTSPRLIKSHLPYRFLPSPDLHNGDSKVI 129
DB 67 VEKCKEDALFNRIPLYECRNEDLINGIKOLAKESPRIVKTHLPKVLPAFWKNCXKI 126
QY 130 YMARNPKOLVSVYQFHRSLRMTSYRGTFQEFRCRPMNDKLGYSWFEHVQEFWEHRMDS 189
DB 127 YLCRNAKDVAVSYYYFLMLITSYPNPKSFSEFVEKFMQGVPGYSWYDHWKAWWKSXNS 186
QY 190 NVLFUKYEDMHRDLVTWVEQLARFLGVSCDKAQLAALTEHCH-----QLVDQ 236
DB 187 RVLFNFIYEDMKDIRREVVKLIIEFLERKPSAELVDRIIQTHTSFQEMKNPNSTNYTMPEE 246
QY 237 CCNAEALP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
DB 247 MNQKVSPFMRKGIIGDMKNHFPPEALRERFDEHYKQMKDCTVKF 291

RESULT 17
US-09-609-816-5
; Sequence 5, Application US/09609816
; Patent No. 6436684
; GENERAL INFORMATION:
; APPLICANT: Woodage, Trevor
; APPLICANT: Wei, Minh Hui
; APPLICANT: Kodira, Chinappa
; APPLICANT: Beasley, Ellen
; APPLICANT: DiFrancesco, Valentina
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669PCT
; CURRENT APPLICATION NUMBER: US/09/609,816
; CURRENT FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 304
; TYPE: PRT
; ORGANISM: HUMAN
US-09-609-816-5

Query Match      27.6%; Score 422; DB 2; Length 304;
Best Local Similarity 32.3%; Pred. No. 1.9e-39;
Matches 97; Conservative 59; Mismatches 124; Indels 20; Gaps 5;

QY 1 MAESEATPTPTGFEF-SKYEFHGVRLPPRCRGMEIEANFPVRPSDVWIVTYPKSGTS 59
DB 1 MAKIEKNAPTWEKKPELNFNMEVDGPTLILSKWEKVKVCNCFQAKPDDLLIATYKSGTT 60
QY 60 LLQEVVLYVLSQAGDPDEIGLWNIIDEOLPVLE--YP---QPGLDIIKELTSPRLIKSHLPY 114
DB 61 WWHEILDMLNDGDVEKCKRAQTLDRAHFALEKPEHKEKPDLEFVLENSPOLIKTHLPS 120
QY 115 RFLPSPDLHNGDSKVIYMARNPKDLVSVYQFHRSLRMTSYRGTFQEFRCRPMNDKLGYS 174
DB 121 HLIPPSIWKENCKIVYVARNPKDCLVSYHFRHSMFMPDPQNLSEEFVEKFMVSGVGS 180
QY 175 WFEHVQEFWEHRMDSNVLFUKYEDMHRDLVTWVEQLARFLGVSCDKAQLAALTEHCHOLV 234
DB 181 WFDHVKGWAAAKDMHRIYLFYEDIKDKPKREIKELKLEKIDISEETLANKIYHTSPDV 240
QY 235 ---DQCNAEALPVG-----RGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
DB 241 MKQPMNTYTTLPTYSIMDHSISPPFMRKMGPDGWNKYFTVAQNEEFDKDYQKMGAGSTLTF 300
```

```
DB 7 EYEVFGFGRGVLMDKRFTKYWEDVEMFLARPDDLVIAITYPKSGTTWISVVYMIYKSGD 66
QY 74 PDEIGLWNIIDEOLPVLEYPOP---GLDIIKELTSPRLIKSHLPYRFLPSPDLHNGDSKVI 129
DB 67 VEKCKEDALFNRIPLYECRNEDLINGIKOLAKESPRIVKTHLPKVLPAFWKNCXKI 126
QY 130 YMARNPKOLVSVYQFHRSLRMTSYRGTFQEFRCRPMNDKLGYSWFEHVQEFWEHRMDS 189
DB 127 YLCRNAKDVAVSYYYFLMLITSYPNPKSFSEFVEKFMQGVPGYSWYDHWKAWWKSXNS 186
QY 190 NVLFUKYEDMHRDLVTWVEQLARFLGVSCDKAQLAALTEHCH-----QLVDQ 236
DB 187 RVLFNFIYEDMKDIRREVVKLIIEFLERKPSAELVDRIIQTHTSFQEMKNPNSTNYTMPEE 246
QY 237 CCNAEALP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
DB 247 MNQKVSPFMRKGIIGDMKNHFPPEALRERFDEHYKQMKDCTVKF 291

RESULT 17
US-09-609-816-5
; Sequence 5, Application US/09609816
; Patent No. 6436684
; GENERAL INFORMATION:
; APPLICANT: Woodage, Trevor
; APPLICANT: Wei, Minh Hui
; APPLICANT: Kodira, Chinappa
; APPLICANT: Beasley, Ellen
; APPLICANT: DiFrancesco, Valentina
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669PCT
; CURRENT APPLICATION NUMBER: US/09/609,816
; CURRENT FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 304
; TYPE: PRT
; ORGANISM: HUMAN
US-09-609-816-5

Query Match      27.6%; Score 422; DB 2; Length 304;
Best Local Similarity 32.3%; Pred. No. 1.9e-39;
Matches 97; Conservative 59; Mismatches 124; Indels 20; Gaps 5;

QY 1 MAESEATPTPTGFEF-SKYEFHGVRLPPRCRGMEIEANFPVRPSDVWIVTYPKSGTS 59
DB 1 MAKIEKNAPTWEKKPELNFNMEVDGPTLILSKWEKVKVCNCFQAKPDDLLIATYKSGTT 60
QY 60 LLQEVVLYVLSQAGDPDEIGLWNIIDEOLPVLE--YP---QPGLDIIKELTSPRLIKSHLPY 114
DB 61 WWHEILDMLNDGDVEKCKRAQTLDRAHFALEKPEHKEKPDLEFVLENSPOLIKTHLPS 120
QY 115 RFLPSPDLHNGDSKVIYMARNPKDLVSVYQFHRSLRMTSYRGTFQEFRCRPMNDKLGYS 174
DB 121 HLIPPSIWKENCKIVYVARNPKDCLVSYHFRHSMFMPDPQNLSEEFVEKFMVSGVGS 180
QY 175 WFEHVQEFWEHRMDSNVLFUKYEDMHRDLVTWVEQLARFLGVSCDKAQLAALTEHCHOLV 234
DB 181 WFDHVKGWAAAKDMHRIYLFYEDIKDKPKREIKELKLEKIDISEETLANKIYHTSPDV 240
QY 235 ---DQCNAEALPVG-----RGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
DB 241 MKQPMNTYTTLPTYSIMDHSISPPFMRKMGPDGWNKYFTVAQNEEFDKDYQKMGAGSTLTF 300
```



```
RESULT 18
US-10-199-334-5
; Sequence 5, Application US/10199334
; Patent No. 6905855
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-3
; CURRENT APPLICATION NUMBER: US/10/199,334
; PRIORITY FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-334-5

Query Match 27.6%; Score 422; DB 2; Length 304;
Best Local Similarity 32.3%; Pred. No. 1.9e-39;
Matches 97; Conservative 59; Mismatches 124; Indels 20; Gaps 5;

QY 1 MAESEAETPSTPGFE-SKYFEFHGVRLLPPFCRGMEBIANFPVRPSDVWIVTPKSGTS 59
Db 1 MAKIEKNAPTMEKPELFNIMEVDGVTLLSKWEKVCNFOAKPDDLILATYPKSGTT 60
QY 60 LLOEVVYLVSQGADPDEIGLNMIDQLPVLE--YP---OPGLDIKELTSPRLIKSHLPY 114
Db 61 WMHEILDMLNDGDVEKCKRAQTLDRAFLFKPHKEKPDLEFVLEMSSPOLIKTHLPS 120
QY 115 RFLPSDLHGDSKVIYMARNPCKDLVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYS 174
Db 121 HLIPIPSWKENCKIVYVARNPKDCLVSYHFRMASFMPDPQNLEEFYEKFMGKVGGS 180
QY 175 WFEHVQEFWEHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLALETCHQLV 234
Db 181 WFDVKGWMAAKDMHRILYLFYEDIKDKPKREIKLKFLKDISEILNKIYHTSPDV 240
QY 235 ---DQCCNAEALPVG-----RGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 241 MKQNPMTNTTLLPTSIMDHSISPFMRKMGPDGKNYFTVAQNEEFDKDYQKMGAGSTLTF 300

RESULT 19
US-10-199-329-5
; Sequence 5, Application US/10199329
; Patent No. 6953681
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-2
; CURRENT APPLICATION NUMBER: US/10/199,329
; PRIORITY FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
US-10-199-329-5

Query Match 27.6%; Score 422; DB 2; Length 304;
Best Local Similarity 32.3%; Pred. No. 1.9e-39;
Matches 97; Conservative 59; Mismatches 124; Indels 20; Gaps 5;

QY 1 MAESEAETPSTPGFE-SKYFEFHGVRLLPPFCRGMEBIANFPVRPSDVWIVTPKSGTS 59
Db 1 MAKIEKNAPTMEKPELFNIMEVDGVTLLSKWEKVCNFOAKPDDLILATYPKSGTT 60
QY 60 LLOEVVYLVSQGADPDEIGLNMIDQLPVLE--YP---OPGLDIKELTSPRLIKSHLPY 114
Db 61 WMHEILDMLNDGDVEKCKRAQTLDRAFLFKPHKEKPDLEFVLEMSSPOLIKTHLPS 120
QY 115 RFLPSDLHGDSKVIYMARNPCKDLVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYS 174
Db 121 HLIPIPSWKENCKIVYVARNPKDCLVSYHFRMASFMPDPQNLEEFYEKFMGKVGGS 180
QY 175 WFEHVQEFWEHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLALETCHQLV 234
Db 181 WFDVKGWMAAKDMHRILYLFYEDIKDKPKREIKLKFLKDISEILNKIYHTSPDV 240
QY 235 ---DQCCNAEALPVG-----RGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 241 MKQNPMTNTTLLPTSIMDHSISPFMRKMGPDGKNYFTVAQNEEFDKDYQKMGAGSTLTF 300
```

```
QY 115 RELPSDLHNGDSKVIYMARNPDKDLVVSYYQFHRSLRTMSYRGTFQFCRRFMDNKLGYGS 174
Db 121 HLIPIPSIWKENCKIVYVAERNPKDCLVSYHYHFRMASFMPDPQNLBEFYEKFMKGKVVYGS 180
QY 175 WFEHVQEFWEHRMDSNVLFLKYEDMHRDLVTWVEQLARFLGVSCDKAQLEALTEHCHQLV 234
Db 181 WFDHVGWGAADKMRHILYLFYEDIKDKPKRIEKLKLEKIDISEILNKIIYHTSFV 240
QY 235 ---DQCNAEALPVG-----RGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTF 280
Db 241 MKQNPWTNYTLPTSIMDHSISPFMRKMGPGDKWKNYFTVAQNEEDKDYQKMGAGSTLTF 300

RESULT 21
US-09-609-816-8
; Sequence 8, Application US/09609816
; Patent No. 6436684
; GENERAL INFORMATION:
; APPLICANT: Woodage, Trevor
; APPLICANT: Wei, Minh Hui
; APPLICANT: Kodira, Chinappa
; APPLICANT: Beasley, Ellen
; APPLICANT: DiFrancesco, Valentina
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; FILE REFERENCE: CL000669PCT
; CURRENT APPLICATION NUMBER: US/09/609,816
; CURRENT FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 283
; TYPE: PRT
; ORGANISM: HUMAN
US-09-609-816-8

Query Match 27.4%; Score 419.5; DB 2; Length 283;
Best Local Similarity 32.6%; Pred. No. 3.3e-39;
Matches 91; Conservative 57; Mismatches 112; Indels 19; Gaps 4;

QY 21 EFHGVLPPFCRGKMBEITANFVRPSDVWIVTYPKSGTSLLOEVVYLVSQGADPDPEIGLM 80
Db 1 EVNGILMSKLSMDNWKIWNFOAKPDDLLIATYAKAGTTWTQEIIVDMIQNDGVDQKQORA 60
QY 81 NIDEOLPVLEY--POP---GLDIKELTSPRLIKSHLPVRLPSPDLHNGDSKVIYMARNP 135
Db 61 NYDRHPFIEWTLPSPLNSGLDLANKMPSRTLKTHLPVHMLPPSFWKENSIIIVARNA 120
QY 136 KDLVVSYYQFHRSLRTMSYRGTFQFCRRFMDNKLGYGSWFHVEHQRMDSNVLFLK 195
Db 121 KCLVSYYYFYSRMNKMPLDPGTLGVEIQFAGKVLGWSWDYHVGWWDVKDQHRILYLF 180
QY 196 YEDMHRDLVTWVEQLARFLGVSCDKAQLEALTEHCHQLV---DQCNAEALP----- 244
Db 181 YEDMKDKPKRIEKKIAKLEKIDISEVLNKIIYHTSFVDMKENPNMANYTTLPSSIMDHSI 240
QY 245 ---VGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTF 280
Db 241 SPFMRKMGPGDKWKNYFTVAQSEDFEDYRRKMGAGSNITF 279

RESULT 22
US-09-609-816-9
; Sequence 9, Application US/09609816
; Patent No. 6436684
; GENERAL INFORMATION:
; APPLICANT: Woodage, Trevor et al.
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; FILE REFERENCE: CL000669D1V-3
; CURRENT APPLICATION NUMBER: US/10/199,334
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; GENERAL INFORMATION:
; APPLICANT: Woodage, Trevor
; APPLICANT: Wei, Minh Hui
; APPLICANT: Kodira, Chinappa
; APPLICANT: Beasley, Ellen
; APPLICANT: DiFrancesco, Valentina
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; FILE REFERENCE: CL000669PCT
; CURRENT APPLICATION NUMBER: US/09/609,816
; CURRENT FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 283
; TYPE: PRT
; ORGANISM: HUMAN
US-09-609-816-9

Query Match 27.4%; Score 419.5; DB 2; Length 283;
Best Local Similarity 32.6%; Pred. No. 3.3e-39;
Matches 91; Conservative 57; Mismatches 112; Indels 19; Gaps 4;

QY 21 EFHGVLPPFCRGKMBEITANFVRPSDVWIVTYPKSGTSLLOEVVYLVSQGADPDPEIGLM 80
Db 1 EVNGILMSKLSMDNWKIWNFOAKPDDLLIATYAKAGTTWTQEIIVDMIQNDGVDQKQORA 60
QY 81 NIDEOLPVLEY--POP---GLDIKELTSPRLIKSHLPVRLPSPDLHNGDSKVIYMARNP 135
Db 61 NYDRHPFIEWTLPSPLNSGLDLANKMPSRTLKTHLPVHMLPPSFWKENSIIIVARNA 120
QY 136 KDLVVSYYQFHRSLRTMSYRGTFQFCRRFMDNKLGYGSWFHVEHQRMDSNVLFLK 195
Db 121 KCLVSYYYFYSRMNKMPLDPGTLGVEIQFAGKVLGWSWDYHVGWWDVKDQHRILYLF 180
QY 196 YEDMHRDLVTWVEQLARFLGVSCDKAQLEALTEHCHQLV---DQCNAEALP----- 244
Db 181 YEDMKDKPKRIEKKIAKLEKIDISEVLNKIIYHTSFVDMKENPNMANYTTLPSSIMDHSI 240
QY 245 ---VGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTF 280
Db 241 SPFMRKMGPGDKWKNYFTVAQSEDFEDYRRKMGAGSNITF 279

RESULT 23
US-10-199-334-8
; Sequence 8, Application US/10199334
; Patent No. 6905855
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; FILE REFERENCE: CL000669D1V-3
; CURRENT APPLICATION NUMBER: US/10/199,334
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
```



```
; FILE REFERENCE: CL000669DIV-2
; CURRENT APPLICATION NUMBER: US/10/199,329
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-329-9

Query Match      27.4%; Score 419.5; DB 2; Length 283;
Best Local Similarity 32.6%; Pred. No. 3.3e-39;
Matches 91; Conservative 57; Mismatches 112; Indels 19; Gaps 4;

QY 21 EFHGVRLPPFCRGKMEIEIANFPVRPSDVWIVYTPKSGTSLQEVVYLVSGADPDDEIGLM 80
DB 1 EVNGILMSKLSMDNDWKIWNFQAKPDDLLIATYAKAGTTWTQEIIVDMIQNDGDVQKQORA 60

QY 81 NIDEQLPVLEY--POP---GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNP 135
DB 61 NYDRHPFIEWTLPSPLNSGLDLANKMPSPTLTKLHPVHMLPPSPFKWENSKIIYVARNA 120

QY 136 KDLVSVYQFHRSLRTMSYRGTFQFCRRFNMNDKLGYSWFHVEHQEFWEHRMDSNVLFK 195
DB 121 KDLVSVYQFHRSLRTMSYRGTFQFCRRFNMNDKLGYSWFHVEHQEFWEHRMDSNVLFK 180

QY 196 YEDMHRDLVTWVEQLARFLVSCDKAQLEALTECHQLV---DQCCNAEALP----- 244
DB 181 YEDMKEDPKREIKKIAKLEKDISEEVNLKIIYHTSPDVMMKENPMANYTTLPSSIMDHSI 240

QY 245 ---VGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTF 280
DB 241 SPFMRKGMGPDWKNYFTVAQSEDFEDYRRKMGAGSNITF 279

RESULT 28
US-10-199-330-9
; Sequence 9, Application US/10199330
; Patent No. 6967094
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; FILE REFERENCE: CL000669DIV-1
; CURRENT APPLICATION NUMBER: US/10/199,330
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-330-9

Query Match      27.4%; Score 419.5; DB 2; Length 283;
Best Local Similarity 32.6%; Pred. No. 3.3e-39;
Matches 91; Conservative 57; Mismatches 112; Indels 19; Gaps 4;

QY 21 EFHGVRLPPFCRGKMEIEIANFPVRPSDVWIVYTPKSGTSLQEVVYLVSGADPDDEIGLM 80
DB 1 EVNGILMSKLSMDNDWKIWNFQAKPDDLLIATYAKAGTTWTQEIIVDMIQNDGDVQKQORA 60

QY 81 NIDEQLPVLEY--POP---GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNP 135
DB 61 NYDRHPFIEWTLPSPLNSGLDLANKMPSPTLTKLHPVHMLPPSPFKWENSKIIYVARNA 120

QY 136 KDLVSVYQFHRSLRTMSYRGTFQFCRRFNMNDKLGYSWFHVEHQEFWEHRMDSNVLFK 195
DB 121 KDLVSVYQFHRSLRTMSYRGTFQFCRRFNMNDKLGYSWFHVEHQEFWEHRMDSNVLFK 180

QY 196 YEDMHRDLVTWVEQLARFLVSCDKAQLEALTECHQLV---DQCCNAEALP----- 244
DB 181 YEDMKEDPKREIKKIAKLEKDISEEVNLKIIYHTSPDVMMKENPMANYTTLPSSIMDHSI 240

QY 245 ---VGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTF 280
DB 241 SPFMRKGMGPDWKNYFTVAQSEDFEDYRRKMGAGSNITF 279

RESULT 27
US-10-199-330-8
; Sequence 8, Application US/10199330
; Patent No. 6967094
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; FILE REFERENCE: CL000669DIV-1
; CURRENT APPLICATION NUMBER: US/10/199,330
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-330-8

Query Match      27.4%; Score 419.5; DB 2; Length 283;
```

[illegible]

```
Db 1 EVNGILSKMSENWDKIWNFOAKPDDLLIATYAKAGTTWTQEIYDVIQNDGDVQKQORA 60
Qy 81 NIDQLPVLVEYPOQ-----GLDIIKELTSPRLIKSHLPYRFLPSDLNNGDSKVYIMARNP 135
Db 61 NTDYRHPPIEWTLPPPLNSGLDLANKMPSRPTLTKLHPVQMLPPSPFKWENSIQIIVVARNA 120
Qy 136 KDLVSVYQFHRSLRTMSYRGTFFQFCRRFMNDKLGYSWFEHVQEFWEHRMDSNVLPFLK 195
Db 121 KDLVSVYQFHRSLRTMSYRGTFFQFCRRFMNDKLGYSWFEHVQEFWEHRMDSNVLPFLK 180
Qy 196 YEDMHRDLVTWVEQLARFLGVSCDKAQLAEALTEHCHQLV---DOCCNAEALP----- 244
Db 181 YEDMKEDPKREIKKIVKLEKIDSEEVLNKIIHHTSPDMQKPNMANYTTLTLPSSIMDHSI 240
Qy 245 ---VGRGVLGWKOIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 241 SPFMRKMGPDGWNKYFTVAQSEDFEDYRKMGAGSTIFF 279

RESULT 32
US-10-199-330-10
; Sequence 10, Application US/10199330
; Patent No. 6967094
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-1
; CURRENT APPLICATION NUMBER: US/10/199,330
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-330-10

Query Match 26.7%; Score 408.5; DB 2; Length 283;
Best Local Similarity 31.5%; Pred. No. 5.9e-38;
Matches 88; Conservative 57; Mismatches 115; Indels 19; Gaps 3;

Qy 21 EFHGVRLPPFCRGKMEETANFPVRPSDVWIVTPYKSGTSLQGVVYVLSQAGADPEIGLM 80
Db 1 EVNGILSKMSENWDKIWNFOAKPDDLLIATYAKAGTTWTQEIYDVIQNDGDVQKQORA 60
Qy 81 NIDQLPVLVEYPOQ-----GLDIIKELTSPRLIKSHLPYRFLPSDLNNGDSKVYIMARNP 135
Db 61 NTDYRHPPIEWTLPPPLNSGLDLANKMPSRPTLTKLHPVQMLPPSPFKWENSIQIIVVARNA 120
Qy 136 KDLVSVYQFHRSLRTMSYRGTFFQFCRRFMNDKLGYSWFEHVQEFWEHRMDSNVLPFLK 195
Db 121 KDLVSVYQFHRSLRTMSYRGTFFQFCRRFMNDKLGYSWFEHVQEFWEHRMDSNVLPFLK 180
Qy 196 YEDMHRDLVTWVEQLARFLGVSCDKAQLAEALTEHCHQLV---DOCCNAEALP----- 244
Db 181 YEDMKEDPKREIKKIVKLEKIDSEEVLNKIIHHTSPDMQKPNMANYTTLTLPSSIMDHSI 240
Qy 245 ---VGRGVLGWKOIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 241 SPFMRKMGPDGWNKYFTVAQSEDFEDYRKMGAGSTIFF 279

RESULT 33
```

```
US-09-609-816-6
; Sequence 6, Application US/09609816
; Patent No. 6436684
; GENERAL INFORMATION:
; APPLICANT: Woodage, Trevor
; APPLICANT: Wei, Minh Hui
; APPLICANT: Kodira, Chinappa
; APPLICANT: Beasley, Ellen
; APPLICANT: DiFrancesco, Valentina
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669PCT
; CURRENT APPLICATION NUMBER: US/09/609,816
; CURRENT FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 304
; TYPE: PRT
; ORGANISM: HUMAN
US-09-609-816-6

Query Match 26.6%; Score 407; DB 2; Length 304;
Best Local Similarity 31.3%; Pred. No. 9.8e-38;
Matches 94; Conservative 61; Mismatches 125; Indels 20; Gaps 5;

Qy 1 MAESEAEPTSPGFE-SKYFEPHGVRLPPFCRGKMEETANFPVRPSDVWIVTPYKSGTS 59
Db 1 MAKTEKNAPTMEKKPELNFMEVDGVPVTLILSKWEKVCNFOAKPDDLILATPKSGTT 60
Qy 60 LLQEVVYVLSQAGADPEIGLMNIDEQLPVLE--YP---QPGLDIIKELTSPRLIKSHLPY 114
Db 61 WMHEILDMILNDGDVEKCKRAQTLDRHAFLELKFPHKEKPDLEFVLEMSPLIKTHLPS 120
Qy 115 RFLPSDLHNGDSKVYIMARNPKDLVSVYQFHRSLRTMSYRGTFFQFCRRFMNDKLGYS 174
Db 121 HLIPPSIWKENCKIVYVARNPKDCLVSYHFRHMASFMPDQNLLEEFYEKEMSGKVGG 180
Qy 175 WFEHVQEFWEHRMDSNVLFKLYEDMHRDLVTWVQLARFLGVSCDKAQLAEALTEHCHQLV 234
Db 181 WFDHVGWAAKQDTHRIYLYFYEDIKNPKRHEIHKVLEFLEKTLSGDVINKIVHHTSFDV 240
Qy 235 ---DOCCNAEALP-----VGRGVLGWKOIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 241 MKONPMANHTAVPAHIFNHSISKFMKMGPDGMKGNHFTVAMNENFDKHVEKMGAGSTLNF 300

RESULT 34
US-10-199-334-6
; Sequence 6, Application US/10199334
; Patent No. 6905855
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-3
; CURRENT APPLICATION NUMBER: US/10/199,334
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
```



```
; APPLICANT: Wei, Minh Hui
; APPLICANT: Kodira, Chinappa
; APPLICANT: Beasley, Ellen
; APPLICANT: DiFrancesco, Valentina
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669PCT
; CURRENT APPLICATION NUMBER: US/09/609,816
; CURRENT FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 304
; TYPE: PRT
; ORGANISM: HUMAN
US-09-609-816-4

Query Match      26.2%; Score 401; DB 2; Length 304;
Best Local Similarity 31.0%; Pred. No. 4.8e-37;
Matches 93; Conservative 62; Mismatches 125; Indels 20; Gaps 5;

Qy 1 MAESEAETPTSGEFE-SKYFEFHGVRLLPPFCRGMEEIANFPVRPSDVWIVTPKSGTS 59
Db 1 MAKIEKNAPTMEKKPELFNIMEVDGPTLLISKWEKVCNFOAKPDDLLIATPKSGTT 60
Qy 60 LLQEVVYLVSOGADPDEIGLNMIDEQLPVLE--YP---OPGLDIIKELTSPRLIKSHLPY 114
Db 61 WMHEILDMLNDGDVEKCKRAQTLDRHAFLELKFPHKEKPDLEFVLEMSSPQLIKTHLPS 120
Qy 115 RFLPSDLHNGDSKVIYMARNPDLVSVYQFHRSLRTMSYRGTFQFCRRFMNDKLGVGYS 174
Db 121 HLIPPSIWKENCKIVYVARNPKDCLVSYHHRMASFMPDPQNLEEFYEKFMGKVGSGS 180
Qy 175 WFEHVQEFWEHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLEALTEHCHQLV 234
Db 181 WFDHVKGWAAAKMHRILYLFYEDIKKPKHEIHKVLEFLEKTSWGDVINKIVHHTSFDV 240
Qy 235 ---DOCCNAEALP-----VGRGRVGLWKDIFTVSMNEKPDLYVKQMGKCDLTFF 280
Db 241 MKDNPMAHTAVPAHIFNHSISKFMKGMGMDGNKHNFTVALNENFDKHYEKKMGAGSTLNF 300

RESULT 38
US-10-199-334-4
; Sequence 4, Application US/10199334
; Patent No. 6905855
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; FILE REFERENCE: CL000669DIV-3
; CURRENT APPLICATION NUMBER: US/10/199,334
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-334-4

Query Match      26.2%; Score 401; DB 2; Length 304;
Best Local Similarity 31.0%; Pred. No. 4.8e-37;
Matches 93; Conservative 62; Mismatches 125; Indels 20; Gaps 5;

Qy 1 MAESEAETPTSGEFE-SKYFEFHGVRLLPPFCRGMEEIANFPVRPSDVWIVTPKSGTS 59
Db 1 MAKIEKNAPTMEKKPELFNIMEVDGPTLLISKWEKVCNFOAKPDDLLIATPKSGTT 60
Qy 60 LLQEVVYLVSOGADPDEIGLNMIDEQLPVLE--YP---OPGLDIIKELTSPRLIKSHLPY 114
Db 61 WMHEILDMLNDGDVEKCKRAQTLDRHAFLELKFPHKEKPDLEFVLEMSSPQLIKTHLPS 120
Qy 115 RFLPSDLHNGDSKVIYMARNPDLVSVYQFHRSLRTMSYRGTFQFCRRFMNDKLGVGYS 174
Db 121 HLIPPSIWKENCKIVYVARNPKDCLVSYHHRMASFMPDPQNLEEFYEKFMGKVGSGS 180
Qy 175 WFEHVQEFWEHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLEALTEHCHQLV 234
Db 181 WFDHVKGWAAAKMHRILYLFYEDIKKPKHEIHKVLEFLEKTSWGDVINKIVHHTSFDV 240
Qy 235 ---DOCCNAEALP-----VGRGRVGLWKDIFTVSMNEKPDLYVKQMGKCDLTFF 280
Db 241 MKDNPMAHTAVPAHIFNHSISKFMKGMGMDGNKHNFTVALNENFDKHYEKKMGAGSTLNF 300

RESULT 38
US-10-199-334-4
; Sequence 4, Application US/10199334
; Patent No. 6905855
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; FILE REFERENCE: CL000669DIV-3
; CURRENT APPLICATION NUMBER: US/10/199,334
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 304
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-334-4

Query Match      26.2%; Score 401; DB 2; Length 304;
Best Local Similarity 31.0%; Pred. No. 4.8e-37;
Matches 93; Conservative 62; Mismatches 125; Indels 20; Gaps 5;

Qy 1 MAESEAETPTSGEFE-SKYFEFHGVRLLPPFCRGMEEIANFPVRPSDVWIVTPKSGTS 59
Db 1 MAKIEKNAPTMEKKPELFNIMEVDGPTLLISKWEKVCNFOAKPDDLLIATPKSGTT 60
Qy 60 LLQEVVYLVSOGADPDEIGLNMIDEQLPVLE--YP---OPGLDIIKELTSPRLIKSHLPY 114
Db 61 WMHEILDMLNDGDVEKCKRAQTLDRHAFLELKFPHKEKPDLEFVLEMSSPQLIKTHLPS 120
Qy 115 RFLPSDLHNGDSKVIYMARNPDLVSVYQFHRSLRTMSYRGTFQFCRRFMNDKLGVGYS 174
Db 121 HLIPPSIWKENCKIVYVARNPKDCLVSYHHRMASFMPDPQNLEEFYEKFMGKVGSGS 180
Qy 175 WFEHVQEFWEHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLEALTEHCHQLV 234
Db 181 WFDHVKGWAAAKMHRILYLFYEDIKKPKHEIHKVLEFLEKTSWGDVINKIVHHTSFDV 240
Qy 235 ---DOCCNAEALP-----VGRGRVGLWKDIFTVSMNEKPDLYVKQMGKCDLTFF 280
Db 241 MKDNPMAHTAVPAHIFNHSISKFMKGMGMDGNKHNFTVALNENFDKHYEKKMGAGSTLNF 300

RESULT 39
US-10-199-329-4
; Sequence 4, Application US/10199329
; Patent No. 6953681
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; FILE REFERENCE: CL000669DIV-2
; CURRENT APPLICATION NUMBER: US/10/199,329
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-329-4

Query Match      26.2%; Score 401; DB 2; Length 304;
Best Local Similarity 31.0%; Pred. No. 4.8e-37;
Matches 93; Conservative 62; Mismatches 125; Indels 20; Gaps 5;

Qy 1 MAESEAETPTSGEFE-SKYFEFHGVRLLPPFCRGMEEIANFPVRPSDVWIVTPKSGTS 59
Db 1 MAKIEKNAPTMEKKPELFNIMEVDGPTLLISKWEKVCNFOAKPDDLLIATPKSGTT 60
Qy 60 LLQEVVYLVSOGADPDEIGLNMIDEQLPVLE--YP---OPGLDIIKELTSPRLIKSHLPY 114
Db 61 WMHEILDMLNDGDVEKCKRAQTLDRHAFLELKFPHKEKPDLEFVLEMSSPQLIKTHLPS 120
Qy 115 RFLPSDLHNGDSKVIYMARNPDLVSVYQFHRSLRTMSYRGTFQFCRRFMNDKLGVGYS 174
Db 121 HLIPPSIWKENCKIVYVARNPKDCLVSYHHRMASFMPDPQNLEEFYEKFMGKVGSGS 180
```

```
Qy 175 WFEHVQEFWEHRMDSNVFLKYEYDMHRDLVTMVEQLARFLGVSCDKAOLEALTEHCHOLV 234
Db 181 WFDHVKGWMAAKDMHRILYLFYEDIKKPKKEIHKVLEFLEKTSWGDVINKIVHHTSDV 240
Qy 235 ---DOCCNAEALP-----VGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTF 280
Db 241 MKDNPMAHNTAVPAHIFNHSISKFMKGMPCGDKNHTVALNENFDKHYEKKMAGSTLNF 300

RESULT 40
US-10-199-330-4
; Sequence 4, Application US/10199330
; Patent No. 6967094
; ORGANISM: Human
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-1
; CURRENT APPLICATION NUMBER: US/10/199,330
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-330-4
```

```
Query Match 26.2%; Score 401; DB 2; Length 304;
Best Local Similarity 31.0%; Pred. No. 4.8e-37;
Matches 93; Conservative 62; Mismatches 125; Indels 20; Gaps 5;

Qy 1 MAESAETPSPGPEPE-SKYEFHGVRLPPFCRCGMKEBIANFPVPRPSDVMTVTPKSGTS 59
Db 1 MAKTEKNAPTWEKKPELFENIMEVDGPTLLLSKEWKEVCNQAKPDDLLIATPKSGTT 60
Qy 60 LLQEVVYLVSQGADPDEIGLMNIDQLPVLE--YP---OPGLDIIKELTSPRLIKSHLPY 114
Db 61 WHHEITDMLNDGDVEKCKRAQTLDRHAFLELKPPEKPKDLEFVLEMSPLQIKTHLPS 120
Qy 115 RFLPSDLHNGDSKVIYMARNPDLVYVYQPHRSRTMSYRGTFQFCRRFMDKLGYS 174
Db 121 HLIPSIWKENCKIVYVARNPKDCLVSYHFRMASFMPDPQNLSEEFYEFKFMGKVGG 180
Qy 175 WFEHVQEFWEHRMDSNVFLKYEYDMHRDLVTMVEQLARFLGVSCDKAOLEALTEHCHOLV 234
Db 181 WFDHVKGWMAAKDMHRILYLFYEDIKKPKKEIHKVLEFLEKTSWGDVINKIVHHTSDV 240
Qy 235 ---DOCCNAEALP-----VGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTF 280
Db 241 MKDNPMAHNTAVPAHIFNHSISKFMKGMPCGDKNHTVALNENFDKHYEKKMAGSTLNF 300
```

```
RESULT 41
US-09-949-016-6526
; Sequence 6526, Application US/09949016
; Patent No. 6812339
; ORGANISM: Human
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
```

```
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6526
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6526

Query Match 24.9%; Score 381; DB 2; Length 350;
Best Local Similarity 33.3%; Pred. No. 1.1e-34;
Matches 94; Conservative 52; Mismatches 108; Indels 28; Gaps 7;

Qy 12 PGEFESKYFEFHGVRLPPFCRG--KMBEIA---NFPVRPSDVMTVTPKSGTSLLOEVV 65
Db 13 PGE-----YFRYKGV---PPFVGLYSLESISLAENTQDVRDDDIIFIITPKSGTTMMIEII 65
Qy 66 YLVSQGADPDEIGLMNIDEQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSDLHNGD 125
Db 66 CLILKEGDPWSWIRSVPIWERAPWCE-TIVGAFSLPDQYSPRLMSSHLPIQIFTKAFFSSK 124
Qy 126 SKVIYMARNPDLVYVYQPHRSRTMSYRGTFQFCRRFMDKLGYSWEHVOEFWEH 185
Db 125 AKVIYMGRRNPDRVVVSLYSHYSKIAGQLKDPQTPQDFLRDLKGEVQFGSWFDHIKGMRLM 184
Qy 186 RMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAOLEALTEH---CHQLVQCCNAEA 242
Db 185 KGDNFLFITTYEELQDQLQSGVERICGFLGRPLGRKEALGVSVAHSTFSAMKANTMSNTYL 244
Qy 243 LP-----VGRGRVGLWKDIFTVSMNEKFDLVYKQK 273
Db 245 LPPSLLDHRRGAFRLKGVCGDKNHTVAQSEAFDRAYRKQM 286
```

```
RESULT 42
US-09-949-016-11211
; Sequence 11211, Application US/09949016
; Patent No. 6812339
; ORGANISM: HUMAN
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11211
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11211
```

```
Query Match 24.9%; Score 381; DB 2; Length 377;
Best Local Similarity 33.3%; Pred. No. 1.3e-34;
Matches 94; Conservative 52; Mismatches 108; Indels 28; Gaps 7;

Qy 12 PGEFESKYFEFHGVRLPPFCRG--KMBEIA---NFPVRPSDVMTVTPKSGTSLLOEVV 65
Db 40 PGE-----YFRYKGV---PPFVGLYSLESISLAENTQDVRDDDIIFIITPKSGTTMMIEII 92
Qy 66 YLVSQGADPDEIGLMNIDEQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSDLHNGD 125
```



```
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; FILE REFERENCE: POLYNUCLEOTIDES ENCODING THE SAME
; CURRENT APPLICATION NUMBER: US/09/795,926
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 303
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-795-926-2

Query Match      19.8%; Score 303.5; DB 2; Length 303;
Best Local Similarity 26.7%; Pred. No. 6.6e-26;
Matches 74; Conservative 61; Mismatches 99; Indels 43; Gaps 9;

QY 19 YFEFHGVRLP-PFCRGK-MEEIANFVRPSDVWIVTPKSGTS-----LLQEVVILVSOQA 72
DB 27 FFTYQIGIPYITMCTSETFQALDTFEARHDDIVLASYPKCGSNWTLHIVSELIYAVSK-- 84
QY 73 DPDEIGLWNIDBQLPVLEYPQ-PGLDI-----IKELTSPRLIKSHLPYRFLPSDLHN 123
DB 85 -----KKYKYPEFPVLECGDSEKQYQRMKGFPSPRILATHLHYDKLPGSIFE 130
QY 124 GDSKVIYMARPKDLVSYQYFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVOBFW 183
DB 131 NKAKILVIFRNPKDTAVSFLHFHNDVDPDIPSYGSMDEFFRQFMKGQVSWGRYDFPAINWN 190
QY 184 EHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAEALTEHCHQLVDQCCNAEL 243
DB 191 KHLGDGNVKFILDYEDLKENLAAGIKQIAEFLGFLTGEOITISV---QSTFOAMRAKSQ 247
QY 244 ----PVG-----RGRVGLWKDIFTVSMNEKFDLVYKQ 271
DB 248 DTHGAVGPFPRKGEVGDWKNLFSIQNQEMDEKFE 284
```

```
RESULT 46
US-10-364-774-2
; Sequence 2, Application US/10364774
; Patent No. 629937
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; FILE REFERENCE: POLYNUCLEOTIDES ENCODING THE SAME
; CURRENT APPLICATION NUMBER: US/10/364,774
```

```
; CURRENT FILING DATE: 2003-02-11
; PRIOR APPLICATION NUMBER: US/09/795,926
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 303
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-364-774-2

Query Match      19.8%; Score 303.5; DB 2; Length 303;
Best Local Similarity 26.7%; Pred. No. 6.6e-26;
Matches 74; Conservative 61; Mismatches 99; Indels 43; Gaps 9;

QY 19 YFEFHGVRLP-PFCRGK-MEEIANFVRPSDVWIVTPKSGTS-----LLQEVVILVSOQA 72
DB 27 FFTYQIGIPYITMCTSETFQALDTFEARHDDIVLASYPKCGSNWTLHIVSELIYAVSK-- 84
QY 73 DPDEIGLWNIDBQLPVLEYPQ-PGLDI-----IKELTSPRLIKSHLPYRFLPSDLHN 123
DB 85 -----KKYKYPEFPVLECGDSEKQYQRMKGFPSPRILATHLHYDKLPGSIFE 130
QY 124 GDSKVIYMARPKDLVSYQYFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVOBFW 183
DB 131 NKAKILVIFRNPKDTAVSFLHFHNDVDPDIPSYGSMDEFFRQFMKGQVSWGRYDFPAINWN 190
QY 184 EHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAEALTEHCHQLVDQCCNAEL 243
DB 191 KHLGDGNVKFILDYEDLKENLAAGIKQIAEFLGFLTGEOITISV---QSTFOAMRAKSQ 247
QY 244 ----PVG-----RGRVGLWKDIFTVSMNEKFDLVYKQ 271
DB 248 DTHGAVGPFPRKGEVGDWKNLFSIQNQEMDEKFE 284
```

```
RESULT 47
US-09-795-926-6
; Sequence 6, Application US/09795926
; Patent No. 6555669
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; FILE REFERENCE: POLYNUCLEOTIDES ENCODING THE SAME
; CURRENT APPLICATION NUMBER: US/09/795,926
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
```



```
; TITLE OF INVENTION: TRANSGENIC PLANTS INCORPORATING TRAIT OF ZOSTERA MARINA
; FILE REFERENCE: PHA-007.01
; CURRENT APPLICATION NUMBER: US/09/854,122
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/202,529
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 324
; TYPE: PRT
; ORGANISM: Brassica napus
US-09-854-122-18

Query Match 17.4%; Score 265.5; DB 2; Length 324;
Best Local Similarity 26.8%; Pred. No. 1.6e-21;
Matches 80; Conservative 61; Mismatches 128; Indels 29; Gaps 11;

QY 9 PSTGFEFSKYFEFHG-VRLPPFCRGMEEIANFPVRPSDVWIVTPKSGTSLLOEVVY- 66
DB 28 PSEKWLVCQVQFQGRWHTQALLQILTCQHFKAQSDIILVTNPKSGTTWLKALVFA 87
QY 67 LVSOQADPEIGLNMIDQLPVLVEYPOPLDIK-----ELTSPLRIKSHLPVRFPLPS 119
DB 88 LINRHKFP-VYSVIIISCVQSALLVPFLGRSLRSPDFDFSQSSPRLMNTHTSHLSLPE 146
QY 120 DLHNGDSKVIYARNPKDLVSYVQFHRSL----RTMSYRGTFQFECRRFMDKLGYSWFW 176
DB 147 SVKSSCKIVYCCRNPKDMFVSLWHFGKLAPEETADY--PIEKAVEAFCEGKFGPW 204
QY 177 EHVQFWEHRMD--SNVLFLKYEDMHRDLVTMVEQLARFLGVS-CDKAQLEALTEHC--H 231
DB 205 DHVLEYWASLSPNPKNLVFSVYEEPKKTGETIKRIAEFLGGLGVEEVRVAILKCSFE 264
QY 232 QLVDOCCNAEA-LPVG-----RGRVGLMKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
DB 265 SLSSLEVNREGKLPSCMETRAFFRKGEGVGRDITLESIAEVIDRTIEBKFGSLKF 322

RESULT 51
US-09-854-122-19
; Sequence 19, Application US/09854122
; Patent No. 6841718
; GENERAL INFORMATION:
; APPLICANT: ALBERTE, RANDALL S.
; APPLICANT: SMITH, ROBERT
; TITLE OF INVENTION: TRANSGENIC PLANTS INCORPORATING TRAIT OF ZOSTERA MARINA
; FILE REFERENCE: PHA-007.01
; CURRENT APPLICATION NUMBER: US/09/854,122
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/202,529
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-854-122-19

Query Match 17.0%; Score 259.5; DB 2; Length 302;
Best Local Similarity 26.9%; Pred. No. 7e-21;
Matches 77; Conservative 52; Mismatches 108; Indels 49; Gaps 12;

QY 9 PSTGFEFSKYFEFHG-VRLPPFCRGMEEIANFPVRPSDVWIVTPKSGTSLLOEVVYL 67
DB 28 PREKGLVSEIVEFQGLWHTQAILQILICQKRFKAQSDIILVTNPKSGTTWLKALVFA 87
QY 68 V-----SQADPEIGLNMIDQLPVL-----YPOGLDIIKELTSPLRIKSHLPYR 115
DB 88 LINRHKFPVSSGNHP--LLVTNPHLLVPFLGVYESSPDFD-FSSLPSRLMNTHTSHL 144
QY 116 FLPSDLHNGDSKVIYARNPKDLVSYVQFHRSL----RTMSYRGTFQFECRRFMDKLG 172
```

```
DB 145 SLPSVSSCKIYVCCRNPKDMFVSLWHFGKLAPEETADY--PIEKAVEAFCEGKFIG 202
QY 173 GSWFEHVQEFW--BHRMDSNVFLKYEDMHRDLVTMVEQLARFLGVSCDKAQLAALTEHC 230
DB 203 GPFWDHILEYWIYASRENPNKVLFTYBELKQTEVEMKRIAETL--ECGFIEEBEVRB-- 258
QY 231 HOLVDQCCNAEA-----LPVG-----RGRVGLMKDIF 257
DB 259 ---IVKLCSPESLSNLEVNKEGKLPNGLETKTTPFRKEIGGWRDSF 301

RESULT 52
US-09-795-926-16
; Sequence 16, Application US/09795926
; Patent No. 6555869
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kiese, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; FILE REFERENCE: LEX-0144-USA
; CURRENT APPLICATION NUMBER: US/09/795,926
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 214
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-795-926-16

Query Match 16.6%; Score 254; DB 2; Length 214;
Best Local Similarity 30.2%; Pred. No. 1.8e-20;
Matches 55; Conservative 42; Mismatches 73; Indels 12; Gaps 3;

QY 99 IKELTSPLRIKSHLPYRFLPSDLHNGDSKVIYARNPKDLVSYVQFHRSLRTMSYRGTF 158
DB 17 MKGPPSPRILATHLHYDKLPGSIFENKAKILVFRNPKDTAVSFLHFHNDVPDIPSYGSW 76
QY 159 QFECRRFMDKLGYSWFEHVQEFWEHRMDSNVFLKYEDMHRDLVTMVEQLARFLGVSC 218
DB 77 DEFQRFQMGQVSWGRYDFDFAINWNKHLDDGONVXFILYEDLKENLAAGIKQIAEFLGFFL 136
QY 219 DKAQLEALTEHCQLVDQCCNAEAAL----PVG-----RGRVGLMKDIFTVSMNEKFDLVY 269
DB 137 TGEQIQTISV---QSTFOAMRAKSQDTHGAVGPFPRKGEVGDWKNLPSLQNMDEKPF 193
QY 270 KQ 271
DB 194 KE 195

RESULT 53
US-10-364-774-16
; Sequence 16, Application US/10364774
; Patent No. 6929937
```

```
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING THE SAME
; FILE REFERENCE: LEX-0144-USA
; CURRENT APPLICATION NUMBER: US/10/364,774
; CURRENT FILING DATE: 2003-02-11
; PRIOR APPLICATION NUMBER: US/09/795,926
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 214
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-364-774-16

Query Match      16.6%; Score 254; DB 2; Length 214;
Best Local Similarity 30.2%; Pred. No. 1.8e-20;
Matches 55; Conservative 42; Mismatches 73; Indels 12; Gaps 3;

QY 99 IKELTSPLRIKSHLPYRFLPSDLHNGDSKVIYMANPKDLVSYQYFHRSLRTMSYRGTF 158
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 17 MKGFPSPRILATHLHYDKLPGSIFENKAKILVIFRNPKDTAVSFLHFHNDVPDIPSGSW 76
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
QY 159 OEFCEFRMNDKLGYSWFEHVOEFWEHRMDSNVLFLKYEDMERDLVTWVEQLARILGVSC 218
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 77 DEFFQFMKGQVSGWRYEDFAINWKNHLDGDNVNFILYEDLKENLAAGIKQIAEFLGPF 136
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
QY 219 DKAQLEALTECHQLVDCCNAEAL----PVG-----RGRVGLWKDIFTVSMNERFLDLVY 269
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 137 TGEQIQTSV---QSTFQMRKASQDTHCAVGPFLEFRKEVGDNKLPSEIQNQEWDEKF 193
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
QY 270 KQ 271
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 194 KE 195

RESULT 54
US-09-795-926-12
; Sequence 12, Application US/09795926
; Patent No. 6555669
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
```

```
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING THE SAME
; FILE REFERENCE: LEX-0144-USA
; CURRENT APPLICATION NUMBER: US/09/795,926
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 186
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-795-926-12

Query Match      12.2%; Score 187; DB 2; Length 186;
Best Local Similarity 26.7%; Pred. No. 6.5e-13;
Matches 46; Conservative 35; Mismatches 59; Indels 32; Gaps 7;

QY 19 YFEHGVRLP-PFCRGK-MEEIANPPVRPSPDVIVTPKSGTS-----LLQEVYVLVSQGA 72
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 27 FFTQGIYPITMCTSETFQALDTFEARHDDIVLASYPKCGSNWILHIVSELIVAVSK-- 84
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
QY 73 DPDEIGLMNIDEQLPVLEYPQ-PGLDI-----IKELTSPLRIKSHLPYRFLPSDLHN 123
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 85 -----KKYKYPEFPVLECGDSEKQYQRMKGFPSPRILATHLHYDKLPGSIFE 130
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
QY 124 GDSKVIYMANPKDLVSYQYFHRSLRTMSYRGTFQEFCEFRMNDKLGYSW 175
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 131 NKAKILVIFRNPKDTAVSFLHFHNDVPDIPSGSWDEFFQFMKGQES-GCW 181
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:

RESULT 55
US-10-364-774-12
; Sequence 12, Application US/10364774
; Patent No. 6929937
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING THE SAME
; FILE REFERENCE: LEX-0144-USA
; CURRENT APPLICATION NUMBER: US/10/364,774
; CURRENT FILING DATE: 2003-02-11
; PRIOR APPLICATION NUMBER: US/09/795,926
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 186
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-364-774-12
```



```
Best Local Similarity 27.5%; Pred. No. 8.3e-12;
Matches 42; Conservative 31; Mismatches 50; Indels 30; Gaps 5;

Qy 36 BEIANFVRPSDVWIVTPKSGTS-----LLOEVVVLVSQGADPDDEIGLNMNIDEQLPVLEY 91
Db 8 QALDTFEARHDDIVLASYPKCGSNWILHIVSELIVAVSK-----KKYKY 51

Qy 92 PQ-PGLDI-----IKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPDKDLVVS 142
Db 52 PEPFVLECGDSEKYQRMKGFPSPRILATHLHYDKLPGSIFENKAKILVIFRNPKDTAVSF 111

Qy 143 YQFHRSLRTMSYRGTFQFCFRFNMNDKLGYSW 175
Db 112 LHFHNDVPDIPSYGSWDEFFRQFMKGQS-GCW 143
```

RESULT 59

```
US-10-364-774-8
; Sequence 8, Application US/10364774
; Patent No. 6929937
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; FILE REFERENCE: LEX-0144-USA
; CURRENT APPLICATION NUMBER: US/10/364,774
; CURRENT FILING DATE: 2003-02-11
; PRIOR APPLICATION NUMBER: US/09/795,926
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 148
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-364-774-8
```

```
Query Match 11.5%; Score 176; DB 2; Length 148;
Best Local Similarity 27.5%; Pred. No. 8.3e-12;
Matches 42; Conservative 31; Mismatches 50; Indels 30; Gaps 5;

Qy 36 BEIANFVRPSDVWIVTPKSGTS-----LLOEVVVLVSQGADPDDEIGLNMNIDEQLPVLEY 91
Db 8 QALDTFEARHDDIVLASYPKCGSNWILHIVSELIVAVSK-----KKYKY 51

Qy 92 PQ-PGLDI-----IKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPDKDLVVS 142
Db 52 PEPFVLECGDSEKYQRMKGFPSPRILATHLHYDKLPGSIFENKAKILVIFRNPKDTAVSF 111

Qy 143 YQFHRSLRTMSYRGTFQFCFRFNMNDKLGYSW 175
Db 112 LHFHNDVPDIPSYGSWDEFFRQFMKGQS-GCW 143
```

RESULT 60

```
US-10-094-749-2841
```

```
; Sequence 2841, Application US/10094749
; Patent No. 6979557
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOTUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2841
; LENGTH: 106
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-2841

Query Match 9.8%; Score 150.5; DB 2; Length 106;
Best Local Similarity 37.2%; Pred. No. 4.1e-09;
Matches 32; Conservative 18; Mismatches 33; Indels 3; Gaps 1;

Qy 36 BEIANFVRPSDVWIVTPKSGTSLLOEVVVLVSQGADPDDEIGLNMNIDEQLPVLEY---P 92
Db 6 DKIWFOAKPDDLLIASYPKAGTTWTQEIVDLIQNDGDIKRRASIQLOHPFLEWTRMT 65

Qy 93 QPGLDIKELTSPRLIKSHLPYRFLP 118
Db 66 HCGIDQANTWPSRPTLKLHPVQLLP 91
```

RESULT 61

```
US-09-795-926-18
; Sequence 18, Application US/09795926
; Patent No. 6555869
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Kieke, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; FILE REFERENCE: LEX-0144-USA
; CURRENT APPLICATION NUMBER: US/09/795,926
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
```

; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 97
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-795-926-18

Query Match 9.0%; Score 137.5; DB 2; Length 97;
Best Local Similarity 35.1%; Pred. No. 1.1e-07;
Matches 27; Conservative 16; Mismatches 33; Indels 1; Gaps 1;
Qy 99 IKELTSPLIKSHLPYRFLPSDLHNGDSKVIYMARPKDLVSVYQFHRSLRTWSYRGTG 158
Db 17 MKGFSPRILATHLHYDKLPGSIFENKAKILVIFRNPKDTAVSFLHFNHNDVPDIPSYGSW 76
Qy 159 QFCRRFNMNKLGYGSW 175
Db 77 DEFRQFMKGES-GCW 92

RESULT 62
US-10-364-774-18
; Sequence 18, Application US/10364774
; Patent No. 6929937
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; FILE REFERENCE: POLYNUCLEOTIDES ENCODING THE SAME
; CURRENT APPLICATION NUMBER: US/10/364,774
; CURRENT FILING DATE: 2003-02-11
; PRIOR APPLICATION NUMBER: US/09/795,926
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 97
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-364-774-18

Query Match 9.0%; Score 137.5; DB 2; Length 97;
Best Local Similarity 35.1%; Pred. No. 1.1e-07;
Matches 27; Conservative 16; Mismatches 33; Indels 1; Gaps 1;
Qy 99 IKELTSPLIKSHLPYRFLPSDLHNGDSKVIYMARPKDLVSVYQFHRSLRTWSYRGTG 158
Db 17 MKGFSPRILATHLHYDKLPGSIFENKAKILVIFRNPKDTAVSFLHFNHNDVPDIPSYGSW 76
Qy 159 QFCRRFNMNKLGYGSW 175

Db 77 DEFRQFMKGES-GCW 92
RESULT 63
US-09-513-999C-8076
; Sequence 8076, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.Pm
; SEQ ID NO 8076
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 9
; OTHER INFORMATION: Xaa=Asp or Glu
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 12
; OTHER INFORMATION: Xaa=Glu or Gly
US-09-513-999C-8076

Query Match 7.5%; Score 115; DB 2; Length 104;
Best Local Similarity 31.2%; Pred. No. 4.5e-05;
Matches 24; Conservative 18; Mismatches 35; Indels 0; Gaps 0;
Qy 14 EFESKYFEFHGVRLPPFCRGMEIEIANFPVRPSDVWITVPKSGTSLQLQEVVYLVSGAD 73
Db 6 DYYXKFXEVHGILMYKDFVYWDNVFAQARPDLLVIATYPKSGTTWVSEIVVMYKREGD 65
Qy 74 PDEIGLWNIQDLPLVLE 90
Db 66 VEKCKEDVIFNRIPFLE 82

RESULT 64
US-09-167-681-49
; Sequence 49, Application US/09167681A
; Patent No. 6265561
; GENERAL INFORMATION:
; APPLICANT: Weinshilboum, M.D., Richard M.
; APPLICANT: Raftogianis, Rebecca B.
; APPLICANT: Wood, Thomas C.
; APPLICANT: Otterness, Diane M.
; TITLE OF INVENTION: SULFOTRANSFERASE SEQUENCE VARIANTS
; FILE REFERENCE: 07039/118001
; CURRENT APPLICATION NUMBER: US/09/167,681A
; CURRENT FILING DATE: 1998-10-07
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 49
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-167-681-49

Query Match 7.3%; Score 111; DB 2; Length 42;
Best Local Similarity 51.3%; Pred. No. 3.3e-05;
Matches 20; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

```
Qy 128 VYIARNPKDLVVSYQFHSRLRSTMSYRGTFQFCCRPM 166
Db 1 VYIARNPKDLVVSYQFHSRLRSTMSYRGTFQFCCRPM 166
; PRIORITY: 2002-04-19
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-286-606-25

Query Match 7.0%; Score 106.5; DB 2; Length 325;
Best Local Similarity 23.7%; Pred. No. 0.0024;
Matches 65; Conservative 37; Mismatches 105; Indels 67; Gaps 14;

Qy 43 VRPSDWMIVTPKSGTSLQEVV-YLVSGQAD-PDEIGLMN--IDEQLPVLEYPOPGLDI 98
Db 27 LRGDGIIISAPSKSLGTQRLVSLVFDGDLPGPLSTVSPWLDQTRPIEEVVATLDA 86
Qy 99 IKELTSPLRIKSHLPYRFLPSDLHNGDSKVIY--MARNPKDLVVS--YVQ----- 144
Db 87 QQH---RRFIKTHPTLDGLVL-----DDRSYICVGRDPRDAAVSMLYQSANNMNEWMRI 138
Qy 145 -----FHRSL-----RTMSYRGTFQFCCRPM--DKLGYGSMFHVQ----- 180
Db 139 LHEAVVPHERIAPPFAELGHARSPTTEF-RDWMEGNPQPPGIGFTHLKGIGTLANILH 197
Qy 181 ---EFWEHRMDSNVLFKYEDMHRDLVTMVQLARFLGVSCDK-----AQLEALTE 228
Db 198 QLGTWVVRHLPNVALFYADYQADLAGELLRPARVLGIAATRDARDLAQYATLDAMS 257
Qy 229 HCHQL-----VDQCCNAEALPVGRGRVGLWKDIFT 258
Db 258 RASEIAPNTTDCGIWHSDEFRFRGGSGDWQOFFT 291

RESULT 67
US-10-891-383A-25
; Sequence 25, Application US/10891383A
; GENERAL INFORMATION:
; APPLICANT: Bertozzi, Carolyn
; APPLICANT: Williams, Spencer J.
; APPLICANT: Mougous, Joseph
; TITLE OF INVENTION: Mycobacterial Sulfation Pathway Proteins
; FILE REFERENCE: BERK-012DIV
; CURRENT APPLICATION NUMBER: US/10/891,383A
; PRIOR FILING DATE: 2004-07-13
; PRIOR FILING DATE: 2001-04-20
; PRIOR FILING DATE: 2001-04-20
; PRIOR FILING DATE: 2001-10-26
; PRIOR FILING DATE: 2001-10-26
; PRIOR FILING DATE: 2002-04-19
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 326
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-891-383A-25

Query Match 7.0%; Score 106.5; DB 2; Length 326;
Best Local Similarity 23.7%; Pred. No. 0.0024;
Matches 65; Conservative 37; Mismatches 105; Indels 67; Gaps 14;

Qy 43 VRPSDWMIVTPKSGTSLQEVV-YLVSGQAD-PDEIGLMN--IDEQLPVLEYPOPGLDI 98
Db 28 LRGDGIIISAPSKSLGTQRLVSLVFDGDLPGPLSTVSPWLDQTRPIEEVVATLDA 87
Qy 99 IKELTSPLRIKSHLPYRFLPSDLHNGDSKVIY--MARNPKDLVVS--YVQ----- 144
Db 88 QQH---RRFIKTHPTLDGLVL-----DDRSYICVGRDPRDAAVSMLYQSANNMNEWMRI 139

Qy 128 VYIARNPKDLVVSYQFHSRLRSTMSYRGTFQFCCRPM 166
Db 1 VYIARNPKDLVVSYQFHSRLRSTMSYRGTFQFCCRPM 166
; PRIORITY: 2002-04-19
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-286-606-25

Query Match 7.0%; Score 106.5; DB 2; Length 325;
Best Local Similarity 23.7%; Pred. No. 0.0024;
Matches 65; Conservative 37; Mismatches 105; Indels 67; Gaps 14;

Qy 43 VRPSDWMIVTPKSGTSLQEVV-YLVSGQAD-PDEIGLMN--IDEQLPVLEYPOPGLDI 98
Db 27 LRGDGIIISAPSKSLGTQRLVSLVFDGDLPGPLSTVSPWLDQTRPIEEVVATLDA 86
Qy 99 IKELTSPLRIKSHLPYRFLPSDLHNGDSKVIY--MARNPKDLVVS--YVQ----- 144
Db 87 QQH---RRFIKTHPTLDGLVL-----DDRSYICVGRDPRDAAVSMLYQSANNMNEWMRI 138
Qy 145 -----FHRSL-----RTMSYRGTFQFCCRPM--DKLGYGSMFHVQ----- 180
Db 139 LHEAVVPHERIAPPFAELGHARSPTTEF-RDWMEGNPQPPGIGFTHLKGIGTLANILH 197
Qy 181 ---EFWEHRMDSNVLFKYEDMHRDLVTMVQLARFLGVSCDK-----AQLEALTE 228
Db 198 QLGTWVVRHLPNVALFYADYQADLAGELLRPARVLGIAATRDARDLAQYATLDAMS 257
Qy 229 HCHQL-----VDQCCNAEALPVGRGRVGLWKDIFT 258
Db 258 RASEIAPNTTDCGIWHSDEFRFRGGSGDWQOFFT 291

RESULT 66
US-10-286-606-25
; Sequence 25, Application US/10286606
; Patent No. 6863895
; GENERAL INFORMATION:
; APPLICANT: Bertozzi, Carolyn
; APPLICANT: Williams, Spencer J.
; APPLICANT: Mougous, Joseph
; TITLE OF INVENTION: Mycobacterial Sulfation Pathway Proteins
; FILE REFERENCE: BERK-012CIP
; CURRENT APPLICATION NUMBER: US/10/286,606
; PRIOR FILING DATE: 2002-10-31
; PRIOR FILING DATE: 2001-04-20
; PRIOR FILING DATE: 2001-04-20
; PRIOR FILING DATE: 2001-10-26
; PRIOR FILING DATE: 2001-10-26
; PRIOR FILING DATE: 2002-04-19
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-286-606-25

Query Match 7.0%; Score 106.5; DB 2; Length 325;
Best Local Similarity 23.7%; Pred. No. 0.0024;
Matches 65; Conservative 37; Mismatches 105; Indels 67; Gaps 14;

Qy 43 VRPSDWMIVTPKSGTSLQEVV-YLVSGQAD-PDEIGLMN--IDEQLPVLEYPOPGLDI 98
Db 27 LRGDGIIISAPSKSLGTQRLVSLVFDGDLPGPLSTVSPWLDQTRPIEEVVATLDA 86
Qy 99 IKELTSPLRIKSHLPYRFLPSDLHNGDSKVIY--MARNPKDLVVS--YVQ----- 144
Db 87 QQH---RRFIKTHPTLDGLVL-----DDRSYICVGRDPRDAAVSMLYQSANNMNEWMRI 138
Qy 145 -----FHRSL-----RTMSYRGTFQFCCRPM--DKLGYGSMFHVQ----- 180
Db 139 LHEAVVPHERIAPPFAELGHARSPTTEF-RDWMEGNPQPPGIGFTHLKGIGTLANILH 197
Qy 181 ---EFWEHRMDSNVLFKYEDMHRDLVTMVQLARFLGVSCDK-----AQLEALTE 228
Db 198 QLGTWVVRHLPNVALFYADYQADLAGELLRPARVLGIAATRDARDLAQYATLDAMS 257
Qy 229 HCHQL-----VDQCCNAEALPVGRGRVGLWKDIFT 258
Db 258 RASEIAPNTTDCGIWHSDEFRFRGGSGDWQOFFT 291
```

Qy 145 -----PHRSL-----RTMSYRGTFQFRCRRFMN--DKLGYGSWFHFVQ----- 180
Db 140 LHEAVVPHERIAPPFAELGHARFTEEF-RDWMEGPNQPPPGIGFTHLKIGITLANILH 198
Qy 181 ---EFWEHRMDSNVLFKYEDHMRDLVTMBQLARFLGVSCDK-----AQLEALTE 228
Db 199 QLGTWVRRLPNVALPHYADQADLACELLRPARVLGIAATRRDRDLAQVATLDAMRS 258
Qy 229 HCHQL-----VQCCNAEALPVGRGVRGLWKDIFT 258
Db 259 RASEIAPNTDGIWHSDEFRFRGGSGDMQOQFFT 292

RESULT 68

US-09-795-926-10
; Sequence 10, Application US/09795926
; Patent No. 6555669
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; FILE REFERENCE: LEX-0144-USA
; CURRENT APPLICATION NUMBER: US/09/795,926
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 148
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-795-926-10

Query Match 6.8%; Score 104.5; DB 2; Length 148;
Best Local Similarity 22.6%; Pred. No. 0.0012;
Matches 30; Conservative 28; Mismatches 44; Indels 31; Gaps 6;
Qy 19 YFEFHGVRLP-PFCRGK-MEEIANPPVRPSDVWIVTPYKSGTS-----LLOEVVYLVSOGA 72
Db 27 FPTYQGIPIPTMCTSETFQALDTFEARHDDIVLASYPKCGSNWLHIVSELIVAVSK-- 84
Qy 73 DPDEIGLMNIDEQLPVLEYPQ-PGLDI-----IKELTSPLIKSHLPYRFLPSDLHN 123
Db 85 -----KKYKYPEFPVLECGDSEKQYRMKGFPSPRILATHLHYDKLPGSIFE 130
Qy 124 GDSKVIYMARNP 136
Db 131 NKAKRQHLTMLPR 143

RESULT 69

US-10-364-774-10
; Sequence 10, Application US/10364774
; Patent No. 6929937
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin

; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; FILE REFERENCE: LEX-0144-USA
; CURRENT APPLICATION NUMBER: US/10/364,774
; CURRENT FILING DATE: 2003-02-11
; PRIOR APPLICATION NUMBER: US/09/795,926
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 148
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-364-774-10

Query Match 6.8%; Score 104.5; DB 2; Length 148;
Best Local Similarity 22.6%; Pred. No. 0.0012;
Matches 30; Conservative 28; Mismatches 44; Indels 31; Gaps 6;
Qy 19 YFEFHGVRLP-PFCRGK-MEEIANPPVRPSDVWIVTPYKSGTS-----LLOEVVYLVSOGA 72
Db 27 FPTYQGIPIPTMCTSETFQALDTFEARHDDIVLASYPKCGSNWLHIVSELIVAVSK-- 84
Qy 73 DPDEIGLMNIDEQLPVLEYPQ-PGLDI-----IKELTSPLIKSHLPYRFLPSDLHN 123
Db 85 -----KKYKYPEFPVLECGDSEKQYRMKGFPSPRILATHLHYDKLPGSIFE 130
Qy 124 GDSKVIYMARNP 136
Db 131 NKAKRQHLTMLPR 143

RESULT 70

US-09-125-635-12
; Sequence 12, Application US/09125635
; Patent No. 6562589
; GENERAL INFORMATION:
; APPLICANT: THE UNITED STATES OF AMERICA represented by THE SE
; TITLE OF INVENTION: AIB1, A novel steroid receptor co-activator
; FILE REFERENCE: 49944
; CURRENT APPLICATION NUMBER: US/09/125,635
; CURRENT FILING DATE: 1998-08-21
; PRIOR APPLICATION NUMBER: 60/049,728
; PRIOR FILING DATE: 1997-06-17
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 1402
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-125-635-12

Query Match 6.8%; Score 104.5; DB 2; Length 1402;
Best Local Similarity 21.0%; Pred. No. 0.038;
Matches 49; Conservative 45; Mismatches 104; Indels 35; Gaps 10;
Qy 14 EFESKYFEFHGVRLPPFCRGKMEIANPPVRPSDVWIVTPYKSGTSLLOEVVYLVSOQAD 73


```

; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; FILE REFERENCE: LEX-0144-USA
; CURRENT APPLICATION NUMBER: US/09/795,926
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 110
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-795-926--4

Query Match          6.1%; Score 93.5; DB 2; Length 110;
Best Local Similarity 22.8%; Pred. No. 0.014;
Matches 26; Conservative 24; Mismatches 35; Indels 29; Gaps 4;

QY 36 EEIANPVRPSDVWIVTPKSGTS---LQEEVVYLVSQGADPEIGLWNIDEQLPVLEY 91
DB 8 QALDTFARHDDIVLASYPKCGSNWLHVSELIYAVSK-----KKYKY 51
QY 92 PQD-PGLDI-----IKELTSPRLKSHLPYRFLPSDLHGNSKVIYMARNP 136
DB 52 PEPFVLECGDSEKYQRMKGFPSPRILATHLHYDKLPGSIFENKAKRQHLTMLPR 105

RESULT 82
US-10-364-774--4
; Sequence 4, Application US/10364774
; Patent No. 6929937
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Halbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING THE SAME
; FILE REFERENCE: LEX-0144-USA
; CURRENT APPLICATION NUMBER: US/10/364,774
; CURRENT FILING DATE: 2003-02-11
; PRIOR APPLICATION NUMBER: US/09/795,926
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 110
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-364-774--4

```

Query Match 6.1%; Score 93.5; DB 2; Length 110;
Best Local Similarity 22.8%; Pred. No. 0.014; Mismatches 24; Indels 29; Gaps 4;
Matches 26; Conservative 24; Mismatches 24; Indels 29; Gaps 4;
QY 36 BEIANFVRPSDVWIVTPKSGTS-----LLOEVVYVLSQGDADPEIGLMNIDEQLPVLVEY 91
DB 8 QALOTFEARHDDIVLASYPKCGSNWILHIVSELIVAVSK-----KKYKY 51
QY 92 PQ-PCLOI-----IKELTSPLIKSHLPYRFLPSDLHNGDSKVIYMARNP 136
DB 52 PEPFVLECGDSEKYQRMKGFSPRILATHLHYDKLPGSFENKAKROHITMLPR 105
RESULT 83
US-09-586-106D-69
; Sequence 69, Application US/09586106D
; Patent No. 6720479
; GENERAL INFORMATION:
; APPLICANT: Wright, David A.
; APPLICANT: Voytas, Daniel F.
; TITLE OF INVENTION: PLANT RETROELEMENTS AND METHODS RELATED THERETO
; FILE REFERENCE: P-1065A
; CURRENT APPLICATION NUMBER: US/09/586.106D
; CURRENT FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: 60/087,125
; PRIOR FILING DATE: 1998-05-29
; PRIOR APPLICATION NUMBER: 09/322,478
; PRIOR FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 69
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Avena sativa
US-09-586-106D-69

Query Match 6.1%; Score 93.5; DB 2; Length 254;
Best Local Similarity 22.9%; Pred. No. 0.051; Mismatches 39; Indels 84; Indels 35; Gaps 10;
Matches 47; Conservative 39; Mismatches 39; Indels 84; Indels 35; Gaps 10;
QY 45 PSDVMIVTPKSGTSLLOEVVYVLSQGDADPEIGLMNIDEQLPVLVEYPPQGLDIIKELTS 104
DB 30 PKKGGITVVPNEDNELIPQVVVVYRMC-----IDFRINKVTRKDHYPFLPFDQMLERLS 85
QY 105 PRLIKSHLPYRFLPSDLHNGDSKVIYMARNPDLVVS-----YYQFHRSL 149
DB 86 K---KTH--FCFL--DGHSQFSQIVVKAQDQKTTFTCPYGTDYRRMPFGLCNAPATFQ 138
QY 150 RTMS--YRGTFQFCRRFMDKLGYSWFE---HVQEFWEHRMDSNVLFLKYEDMH---R 201
DB 139 RCMSAIFHGFCEEIIVEFMDDFSQVGTSDNCLHNLDFLQRFETNLVLNWEKCHPMVN 198
QY 202 DLVTMVEQLARFLGVSCDKAQLAL 226
DB 199 EGVILGHKISE-RGIEVDRAKIEAI 222

Query Match 6.0%; Score 92.5; DB 2; Length 883;
Best Local Similarity 20.8%; Pred. No. 0.44; Mismatches 52; Indels 57; Gaps 11;
Matches 52; Conservative 37; Mismatches 52; Indels 57; Gaps 11;
QY 12 PGFESEKYPFHGVRLPPFCRGKMEIATNFPVRPSDVW-----IVTYPKSGT 58
DB 562 PVPLAQKYFELFPQERSPLNQPCDD-----KRHKDIWSKEKTCRDLKPKFLIVGPQKTGT 616
QY 59 SLLOEVVYL-----VSGADP---DEIGLMN-----IDEQLPVLVEY-PQGLDIIKELT 103
DB 617 TAIHFFLSLHPAVTSFSPSTFEEIQFFNSPNYHKGIDWMTDFPVPNSASTDFLEKS 676
QY 104 SPRLIKSHLPYR---FLPSDLHNGDSKVIYMARNPDLVVSYYQFHRSLR---TMSYRGT 157
DB 677 ATYFDSVVPVRGALLPR-----AKIIVLTAPRAYSWYQORAHGDPVALNY--T 728
QY 158 FQEE-----CRRFMDKLGYSWFEHVQEFWEHRMDSNVLFLKYEDMHRLDVTMV 207

PRIOR FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 69
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Avena sativa
US-10-799-870-69

Query Match 6.1%; Score 93.5; DB 2; Length 254;
Best Local Similarity 22.9%; Pred. No. 0.051; Mismatches 47; Conservative 39; Mismatches 47; Indels 84; Indels 35; Gaps 10;
QY 45 PSDVMIVTPKSGTSLLOEVVYVLSQGDADPEIGLMNIDEQLPVLVEYPPQGLDIIKELTS 104
DB 30 PKKGGITVVPNEDNELIPQVVVVYRMC-----IDFRINKVTRKDHYPFLPFDQMLERLS 85
QY 105 PRLIKSHLPYRFLPSDLHNGDSKVIYMARNPDLVVS-----YYQFHRSL 149
DB 86 K---KTH--FCFL--DGHSQFSQIVVKAQDQKTTFTCPYGTDYRRMPFGLCNAPATFQ 138
QY 150 RTMS--YRGTFQFCRRFMDKLGYSWFE---HVQEFWEHRMDSNVLFLKYEDMH---R 201
DB 139 RCMSAIFHGFCEEIIVEFMDDFSQVGTSDNCLHNLDFLQRFETNLVLNWEKCHPMVN 198
QY 202 DLVTMVEQLARFLGVSCDKAQLAL 226
DB 199 EGVILGHKISE-RGIEVDRAKIEAI 222

RESULT 85
US-09-949-016-6398
; Sequence 6398, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949.016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6398
; LENGTH: 883
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6398

Query Match 6.0%; Score 92.5; DB 2; Length 883;
Best Local Similarity 20.8%; Pred. No. 0.44; Mismatches 52; Indels 57; Gaps 11;
Matches 52; Conservative 37; Mismatches 52; Indels 57; Gaps 11;
QY 12 PGFESEKYPFHGVRLPPFCRGKMEIATNFPVRPSDVW-----IVTYPKSGT 58
DB 562 PVPLAQKYFELFPQERSPLNQPCDD-----KRHKDIWSKEKTCRDLKPKFLIVGPQKTGT 616
QY 59 SLLOEVVYL-----VSGADP---DEIGLMN-----IDEQLPVLVEY-PQGLDIIKELT 103
DB 617 TAIHFFLSLHPAVTSFSPSTFEEIQFFNSPNYHKGIDWMTDFPVPNSASTDFLEKS 676
QY 104 SPRLIKSHLPYR---FLPSDLHNGDSKVIYMARNPDLVVSYYQFHRSLR---TMSYRGT 157
DB 677 ATYFDSVVPVRGALLPR-----AKIIVLTAPRAYSWYQORAHGDPVALNY--T 728
QY 158 FQEE-----CRRFMDKLGYSWFEHVQEFWEHRMDSNVLFLKYEDMHRLDVTMV 207

Db 729 FYOVISASOTPLALRSLQNRCLVPGYGYSTHLQRMWLTYPSPGOLLIVDGOELRTNPAASM 788

QY 208 EQLARFLGVS 217
| : : : : :
Db 789 ESIQKFLGIT 798

RESULT 86
US-09-949-016-7255
; Sequence 7255, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7255
; LENGTH: 892
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7255

Query Match 6.0%; Score 92.5; DB 2; Length 892;
Best Local Similarity 20.8%; Pred. No. 0.45;
Matches 52; Conservative 37; Mismatches 104; Indels 57; Gaps 11;

QY 12 PGFEFSKYFEFGVRLPPFCRGHMEIANFPVRPSDVW-----IVTPKSGT 58
| : : : : :
Db 571 PVPQAQYFELFPQERSPLWQPCDD-----KRHKDINSKEKTCRLPKFLVGPQKTGT 625

QY 59 SLLEQVYVL-----VSQADP---DEIGLNV-----IDEQLPVLEYP-QPGLDIIKELT 103
| : : : : :
Db 626 TAIHFPLSLHDAVTSFSPSPSTFEELQFFNSYHKGIDWYMDFFVPVPSNASTDFLFEKS 685

QY 104 SPRLIKSHLPYR---FLPSDLHNGDSKVIYMARNPDKLVVSYVYQFHRSLR---TMSYRGT 157
| : : : : :
Db 686 ATYFDSVVPVRRGAALLPR-----AKIITVLTNPADRAYSVYQORAHGDPVALNY--T 737

QY 158 POFE-----CRRFVNDKLGYSWFEHVOEFWEHRMDSNVLFKYMEDMHRDLVTMV 207
| : : : : :
Db 738 FYOVISASOTPLALRSLQNRCLVPGYGYSTHLQRMWLTYPSPGOLLIVDGOELRTNPAASM 797

QY 208 EQLARFLGVS 217
| : : : : :
Db 798 ESIQKFLGIT 807

RESULT 87
US-09-270-767-41537
; Sequence 41537, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41537
; LENGTH: 121
; TYPE: PRT
; ORGANISM: Drosophila melanogaster

US-09-270-767-41537

Query Match 5.8%; Score 88; DB 2; Length 121;
Best Local Similarity 26.0%; Pred. No. 0.07;
Matches 38; Conservative 20; Mismatches 32; Indels 56; Gaps 9;

QY 78 GLMNIDEQLPVLEYQPGGLDIIKELTSPRLIKSHLPFLPSDLHNGDSKVIYMARNPDK 137
| : : : : :
Db 14 GLGRADEHVALSQYPRPGM-----LPTKHPNSDKPKL---RNLIK- 49

QY 138 LVVSYVQFHRSLRTMSYRGTGTF-----QEFCCRFRMNDKLGYSWFE-HVQEFWEHRMDSN 190
| : : : : :
Db 50 -IMGY-----SURTDIYRTVMVRFAQNFSR-----DWHDDVYGEELYDHRDLDSG 93

QY 191 -----VLFLEKYED---MHRDLVTMV 207
| : : : : :
Db 94 BELNLVPLPQFDDVRQRLRRRLMEVW 119

RESULT 88
US-09-949-016-8271
; Sequence 8271, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8271
; LENGTH: 610
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8271

Query Match 5.8%; Score 88; DB 2; Length 610;
Best Local Similarity 21.7%; Pred. No. 0.82;
Matches 61; Conservative 38; Mismatches 92; Indels 90; Gaps 15;

QY 21 EFHGVRLPPFCRGHMEIANFPVRPSDVWIVTPYKSGTSLLOEVVYVVSQADPDEIGLM 80
| : : : : :
Db 170 EFH-----SKLQEATNFPRLP---FVIPFLKANLPQLQRELLHCARLAKQNPAYL 217

QY 81 NIDEOL-----PV-----LEYQPGLDIIKELTSPRLIK-----SHLP 113
| : : : : :
Db 218 AQHEQLLLDASTTSPVDSSELLDVNENG---KRTPDRTKENGDFRPLHSEHPKRP 273

QY 114 YRFLPSDLHNGDSKVIYMARNP-----PK-----DLVVSY-----YQFHRSL- 149
| : : : : :
Db 274 CTISPGQRYSPNNGLSYQPNGLPHTPPPPQHRYLDDMAIAHHYRDSYRHPSHRDLRDN 333

QY 150 RTMSYRGTGTFQFCRRFRMNDKLGYSWFEHVOEFWEHRMDSNVLFKYMEDMHRDLVTMV 209
| : : : : :
Db 334 RPMGLHGTRQE---EMIDHRLTDREWAE---WKH-----LDHLLNCIMDWVEK 376

QY 210 LARFLGV--SCDKAQLEALTECHQLVDQCCNAEALPVGRG 248
| : : : : :
Db 377 TRRSLTVLRCQEADEELNYWIRYSD-----AEDLKKGGG 413

RESULT 89
US-09-949-016-8270
; Sequence 8270, Application US/09949016
; Patent No. 6812339

```

RESULT 90
US-09-949-016-7288
; Sequence 7288, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7288
; LENGTH: 702
; TYPE: PRT
; ORGANISM: Human
; US-09-949-016-7288

Query Match          5.8%; Score 88; DB 2; Length 702;
Best Local Similarity 21.7%; Pred. No. 1;

```

Db 312 EFH-----SKLQEAENFPLRP---FVIFPLKANLPLLORELLHLCARLAKQNPAYL 359
QY 81 NIDEQL-----PV-----LEYPOGLDIIKELTSRLIK-----SHLP 113
Db 360 AQHEQLLDASTTSPVDSSELLDVNENG---KRTPTDKENGFDREPLHSEHPKRP 415
QY 114 YRFLPSDLHNGSKVIYMARN-----PK-----DLVVS-----YQFHRSL---- 149
Db 416 CTISPGQRYSPNNGUSYQNGLPHPTPPPPQHYRDLDDMAIAHHYRDSYRHPSHRDLDRN 475
QY 150 RTMSYRGTFQFCRCRRFMDKLGYSWFHVFQEFWEHRMDSNVLFLLKYEDMHRDLVTMV 207
Db 476 RPKGLHGTQE---EMIDHRLTDREWABE-----WKH-----LDHLLNCIMDWVK 518
QY 210 LARFLGV---SCDKAQLAETHCHQLVDCCNAELPVG 248
Db 519 TRRSUTVLRRCCQEADEELNYIRYSYD-----AEDLKGGG 555

RESULT 92

US-09-345-473E-12
; Sequence 12, Application US/09345473E
; Patent No. 6558903
; GENERAL INFORMATION:
; APPLICANT: Hodge, Martin
; TITLE OF INVENTION: No. 6558903el Kinases and Uses Thereof
; FILE REFERENCE: 35800/183781
; CURRENT APPLICATION NUMBER: US/09/345,473E
; PRIORITY FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 172
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(172)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-345-473E-12

Query Match 5.7%; Score 87.5; DB 2; Length 172;
Best Local Similarity 28.8%; Pred. No. 0.14; Mismatches 35; Indels 23; Gaps 5;
Matches 30; Conservative 16;
QY 153 SYRGTFQFCR-----RFMNDKLGYSWFHVFQEFWEHRMDSNVLFLLKYEDMHRDLVTMV 207
Db 38 SYRALISAFSLRLTRLDFTCEKIGSG-PFSEVFKV-RHRASQVMAKXNTLSSNRANML 95
QY 208 EQ-----LARFLGVSCDKAQLAETHCH-----QLVD 235
Db 96 KEVQLMNLRLSHPNILRFMGVCVHQGLHALTEYINSGNLEQLLD 139

RESULT 93

US-09-862-027-12
; Sequence 12, Application US/09862027
; Patent No. 6858418
; GENERAL INFORMATION:
; APPLICANT: Hodge, Martin R.
; TITLE OF INVENTION: No. 6858418el Kinases and Uses Thereof
; FILE REFERENCE: 35800/234862
; CURRENT APPLICATION NUMBER: US/09/862,027
; PRIORITY FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 09/345,473
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 172
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:

; NAME/KEY: VARIANT
; LOCATION: (1)...(172)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-862-027-12
Query Match 5.7%; Score 87.5; DB 2; Length 172;
Best Local Similarity 28.8%; Pred. No. 0.14; Mismatches 35; Indels 23; Gaps 5;
Matches 30; Conservative 16;
QY 153 SYRGTFQFCR-----RFMNDKLGYSWFHVFQEFWEHRMDSNVLFLLKYEDMHRDLVTMV 207
Db 38 SYRALISAFSLRLTRLDFTCEKIGSG-PFSEVFKV-RHRASQVMAKXNTLSSNRANML 95
QY 208 EQ-----LARFLGVSCDKAQLAETHCH-----QLVD 235
Db 96 KEVQLMNLRLSHPNILRFMGVCVHQGLHALTEYINSGNLEQLLD 139

RESULT 94

US-08-946-329A-49
; Sequence 49, Application US/08946329A
; Patent No. 6057091
; GENERAL INFORMATION:
; APPLICANT: Beachy, Philip A.
; APPLICANT: Porter, Jeffrey A.
; TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES
; NUMBER OF SEQUENCES: 109
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/946,329A
; FILING DATE: 07-OCT-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/061,323
; FILING DATE: 07-OCT-1996
; APPLICATION NUMBER: 08/729,743
; FILING DATE: 10-JUL-1996
; APPLICATION NUMBER: 08/567,357
; FILING DATE: 04-DEC-1995
; APPLICATION NUMBER: 08/349,498
; FILING DATE: 02-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/140001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 154 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-946-329A-49

Query Match 5.7%; Score 87; DB 2; Length 154;
Best Local Similarity 24.6%; Pred. No. 0.13;
Matches 32; Conservative 22; Mismatches 48; Indels 28; Gaps 6;
QY 154 YRGTFQFCRRFMDKLGYSWFHVFQEFWEHRMDSNVLFLLKYED-MHRDLVTMVQOLAR 212
Db 11 YNGGFKRM-----DELAUGDWQALD-----KNGSQVTFIPVQVWLHRD-----PKQVAD 55

QY 213 FLGVSCDKAQLEALTEHCHQLVQC-----CNAEALPVGRGVLWKDIFTYSMNE 263
| : | : |||| | | | | : | : | : | : | :
Dp 56 FVEETLDNGETFSLTEKHLVFVTCQSPYSEDENINANPVAERNVG--DCFYTAHRK 112

Qy 264 KFDLVYKQKM 273
pb 113 KSOMYORVKV 122

```

RESULT 95
US-09-562-914-49
; Sequence 49, Application US/09562914
; Patent No. 6911528
; GENERAL INFORMATION:
; APPLICANT: Beachy, Philip A.
; Porter, Jeffrey A.
; TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES
; NUMBER OF SEQUENCES: 109
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037

```

Query Match	5.7%	Score 87	DB 2	Length 154
Best Local Similarity	24.6%	Pred. NO. 0.13		
Matches	32	Conservative	22	Mismatches 48
				Indels 28
				Gaps 6

Qy	213	FLGVSCDKAQLEALTEHCHQLVDQC-----CNAEALPVGRGVLWKDIFTVMNE	263
Dq	56	FVEETLDNGCTESTETKHLVFTOCSPYSEGINANPVAERNVG---DCFYTAHRK	112

```

Qy      264 KFDLVYKQKM 273
      | : | :
Db      113 KSQMYORVKV 122

RESULT 96
US-09-248-796A-14484
; Sequence 14484, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:

```

Query Match	5.7%;	Score 86.5;	DB 2;	Length 499;
Best Local Similarity	23.9%;	Pred. No. 0.9;		
Matches	47;	Conservative	26;	Mismatches 71; Indels 53; Gaps 10;
Qy	40	NFVRPSDVM-----IVTPKSGTSL---LQEVVVLVSGQADPDDEIGLWN- 81		
Db	167	NFIHDDSVWSLYSQSTSHDSFKTFYSGDKSGNIKTDLRSVTFNTRTSNDDFFGFSNY 226		
Qy	82	-----IDQLPVLEVQPGLDIIKELTSPRIKSHLPYR---FLPSDLHNGDSKVYIYWARN 134		
Db	227	ENNYIDEKLGI-----STLIAKSTSPILSICHEPEQDTIFAT-----NMESMTRYI--N 273		
Qy	135	PKDLVSVYGYPHRSLRTWTSYRGTFQBFCCRFPMNDKLGVGSWFEHFQFWEHRMDSNVFL 194		
Db	274	PDTCLSQYQLRQCLDYTN-----TNGIGVGN--DNNDELLHLDGGSNP--- 318		
Qy	195	KYEDMHRDLVTWVEQLA 211		
Db	319	PTEDLNSDFYDLISHL 335		

```

RESULT 97
US-09-445-353E-2
; Sequence 2, Application US/09445353E
; Patent No. 6812336
; GENERAL INFORMATION:
; APPLICANT: Rosenfeld, Michael G.
; APPLICANT: Glass, Christopher K.
; APPLICANT: Rose, David W.
; APPLICANT: Torchia, Joseph
; TITLE OF INVENTION: A Transcription Factor Coactivator Protein, p/CIP
; FILE REFERENCE: 6627-PAL021
; CURRENT APPLICATION NUMBER: US/09/445,353E
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: PCT/US98/12263
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/049,452
; PRIOR FILING DATE: 1997-06-12
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1402
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature

```

1 MAESEAE-----TPSTPGEFFESKYFFEFHG-----VRLPPFCRGKM-----EEIANF 41

[illegible]

Search completed: May 17, 2006, 11:06:27
Job time : 56 secs

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	1530	34.8	284	3	AA79215	Human tra	
2	1530	34.8	284	5	AAU78268	Human ner	
3	1530	34.8	284	6	ABR01809	Human cer	
4	1530	34.8	284	8	ADQ89158	Human uro	
5	1530	34.8	284	8	ADR44884	Polypepti	
6	1530	34.8	284	9	ADZ21933	Human SUL	
7	1506	34.2	284	6	ABM04807	Rat sulfo	
8	1142.5	26.0	218	6	ABR01860	Human can	

QY 501 TTCTCCGGAGGTTTATGAATGATAAGCTGGGCTACGGCTCCTCGTTTGACACGTCGAG 560
 |||||
 Db 161 PheCysArgArgPheMetAsnAspLysLeuGlyTyrGlySerTrpPheGluHisValGln 180
 |||||
 QY 561 GAGTTCTGGGAGCACCAGTCAGTCCGAACGTGCTTTTCTCAAGTATGAGACATGCAT 620
 |||||
 Db 181 GluPheTrpGluHisArgMetAspSerAsnValLeuPheLeuLysTyrGluAspMetHis 200
 |||||
 QY 621 CGGACCTGCTGACGATGGTGGAGCAGCTGCCAGATTCTCTGGGGGTCTCTGTGACAAG 680
 |||||
 Db 201 ArgAspLeuValThrMetValGluGlnLeuAlaArgPheLeuGlyValSerCysAspLys 220
 |||||
 QY 681 GCCAGCTGGAAGCCCTGACCGAGACACTGCCACAGCTGGTGGACAGTGTGCAACGCT 740
 |||||
 Db 221 AlaGlnLeuGluAlaLeuThrGluHisCysHisGlnLeuValAspGlnCysCysAsnAla 240
 |||||
 QY 741 GAGCCCTGCCCGTGGGCGGGGAAGAGTGTGGGTGTGGAGGACATCTTCACCGTCTCC 800
 |||||
 Db 241 GluAlaLeuProValGlyArgGlyValGlyLeuTrpLysAspIlePheThrValSer 260
 |||||
 QY 801 ATGAATGAGAAGTTTGACTGGTGTATATAACAGAAAGATGGAAAGTGTGACCTCACGTTT 860
 |||||
 Db 261 MetAsnGluLysPheAspLeuValTyrLysGlnLysMetGlyLysCysAspLeuThrPhe 280
 |||||
 QY 861 GACTTTATTATTA 872
 |||||
 Db 281 AspPheTyrLeu 284
 |||||

RESULT 2

AAU78268
 ID AAU78268 standard; protein; 284 AA.

AAU78268;

05-JUN-2002 (first entry)

Human nervous system-specific sulfotransferase, SULTn, protein sequence.

Human; nervous system-specific sulfotransferase; SULTn; sulfonation;
 neurological disorder; multiple sclerosis; Huntington's chorea;
 nervous system-mediated muscular atrophy; amyotrophic lateral sclerosis;
 myasthenia gravis; Alzheimer's disease; senile dementia; Schizophrenia;
 bipolar disorder; cytosolic sulfotransferase gene family;
 cerebellar ataxia; chromosome 22q13.

Homo sapiens.

Key Location/Qualifiers
 Binding-site 54..61
 /label= PAPS_5'_phosphosulfate_binding_site
 Binding-site 129..144
 /label= PAPS_3'_phosphate_binding_site
 Binding-site 246..258
 /label= PAPS_3'_phosphate_binding_site

W0200218541-A2.

07-MAR-2002.

04-SEP-2001; 2001WO-US041986.

01-SEP-2000; 2000US-0229929P.

(UYBO-) UNIV BOSTON.

Farb DH, Martin S;

WPI; 2002-281063/32.

N-PSDB; ABK12289.

New nervous system-specific sulfotransferase (SULTn) polypeptide for

treating neurological disorders including multiple sclerosis, nervous

system-mediated muscular atrophy and Huntington's chorea.

Claim 2; Fig 1a; 46pp; English.

The present invention relates to a new nervous system-specific
 sulfotransferase (SULTn) polypeptide. The SULTn polypeptide of the
 invention comprises a nervous system-specific sulfotransferase,
 comprising a sequence of 284 amino acids, given in the specification. The
 invention is useful for identifying an effector of nervous system-
 specific sulfonation by contacting the polypeptide with a candidate
 effector compound and assaying the ability of the candidate effector
 compound to effect the sulfotransferase function of the polypeptide. The
 invention is also useful for inhibiting nervous system-specific
 sulfonation by contacting the polypeptide under physiological conditions
 with a compound capable of interacting with and inhibiting the
 sulfonation activity of the polypeptide, where the compound is an
 antibody, peptide, polypeptide, nucleic acid, organic molecule or an
 inorganic molecule. The polypeptide of the invention is also useful for
 treating a neurological disorder, including multiple lateral sclerosis,
 Huntington's chorea, myasthenia gravis, Alzheimer's disease, senile
 dementia, Schizophrenia, bipolar disorder and cerebellar ataxia. The
 present amino acid sequence represents the human SULTn protein of the
 invention. This sequence is encoded by the human SULTn gene located on
 chromosome 22q13. SULTn is a member of the cytosolic sulfotransferase
 gene family

SQ Sequence 284 AA;

Alignment Scores:

Pred. No.: 3.01e-136 Length: 284
 Score: 1530.00 Matches: 284
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 34.8% Indels: 0
 DB: 5 Gaps: 0

US-10-768-158-1 (1-2419) x AAU78268 (1-284)

QY 21 ATGCGGAGAGCGAGGCGGAGACCCCGAGACCCCGGGGAGTTCGAGAGCAAGTACTTC 80
 |||||
 Db 1 MetAlaGluSerGluAlaGluThrProSerThrProGlyGluPheGluSerLysTyrPhe 20
 |||||
 QY 81 GAGTTCATGCGTGGCGGTGCGCGCTTCTGCGGGGAAGATGGAGGAGATCGCAAC 140
 |||||
 Db 21 GluPheHisGlyValArgLeuProProPheCysArgGlyLysMetGluGluIleAlaAsn 40
 |||||
 QY 141 TTCCTGGTGGCGCCGAGCGAGTGTGATCGTCACTACCTACCCAGTCCCGCAGCGTTG 200
 |||||
 Db 41 PheProValArgProSerAspValTrpIleValThrTyrProLysSerGlyThrSerLeu 60
 |||||
 QY 201 CTGCAGGAGGTGCTACTTTGGTGAGCCAGGGCGCTGACCCCGATGAGATCGGCTTGATG 260
 |||||
 Db 61 LeuGlnGluValValTyrLeuValSerGlnGlyAlaAspProAspGluIleGlyLeuMet 80
 |||||
 QY 261 AACATGACGAGCAGCTCCCGGTCTGTGAGTACCCACAGCCGGCGCTCGGACATCATCAAG 320
 |||||
 Db 81 AsnIleAspGluGlnLeuProValLeuGluTyrProGlnProGlyLeuAspIleIleLys 100
 |||||
 QY 321 GAAGTACCTCTCCCGCTCATCAGAGCCACCTGCGCTACCGCTTCTCGCCTCTGAC 380
 |||||
 Db 101 GluLeuThrSerProArgLeuIleLysSerHisLeuProTyrArgPheLeuProSerAsp 120
 |||||
 QY 381 CTCACAAATGGAGACTCCAAGGTCACTATATGGCTCGCAACCCCAAGGATCTGGTGGTG 440
 |||||
 Db 121 LeuHisAsnGlyAspSerLysValIleTyrMetAlaArgAsnProLysAspLeuValVal 140
 |||||
 QY 441 TCTTATTATCAGTTCACCGCTCTCTCGGACCATGAGTACCGAGGCACCTTTCAAGAA 500
 |||||
 Db 141 SerTyrTyrGlnPheHisArgSerLeuArgTyrMetSerTyrArgGlyThrPheGlnGlu 160
 |||||
 QY 501 TTCTCCGGAGGTTTATGAATGATAAGCTGGGCTACGGCTCCTCGTTTGACACGTCGAG 560
 |||||

Db 161 PheCysArgArgPheMetAsnAspLysLeuGlyTyrGlySerTrpPheGluHisValGln 180
QY 561 GAGTTCGGGAGCACCAGCTCGAAGCTGCTTTTCTCRAAGTATGAAGACATGCAT 620
Db 181 GluPheTrpGluHisArgMetAspSerAsnValLeuPheLeuYsYrGluAspMetHis 200
QY 621 CGGACCTGGTGGAGCTGGAGCAGCTGCCAGATTCTTGGGGGTGCTCTGTGACAAG 680
Db 201 ArgAspLeuValThrMetValGluGlnLeuAlaArgPheLeuGlyValSerCysAspLys 220
QY 681 GCCAGCTGGAAGCCCTGACGAGACATGCCACAGCTGGTGGACCAAGTGTGCAACGCT 740
Db 221 AlaGlnLeuGluAlaLeuThrGluHisCysHisGlnLeuValAspGlnCysCysAsnAla 240
QY 741 GAGCCCTGCCGTCGGCGGGAGAGTGGGCTGTGGAGGACATCTTCACCGTCTCC 800
Db 241 GluAlaLeuProValGlyArgGlyArgValGlyLeuTrpLysAspIlePheThrValSer 260
QY 801 ATGAATGAGAAGTTGACTTGGTGTATATAACAGAGAAGATGGAAAGTGTGACCTCACGTTT 860
Db 261 MetAsnGluLysPheAspLeuValTyrLysGlnLysMetGlyLysCysAspLeuThrPhe 280
QY 861 GACTTTTATTTA 872
Db 281 AspPheTyrLeu 284

RESULT 3

ABR01809
ID ABR01809 standard; protein; 284 AA.
XX AC ABR01809;
XX 19-MAY-2003 (first entry)
DT Human cancer-related protein, 192P2G7.
DE Human; cytostatic; vaccine; cancer; immune response.
KW Homo sapiens.
OS WO200283921-A2.
XX 24-OCT-2002.
XX 10-APR-2002; 2002WO-US011654.
XX 10-APR-2001; 2001US-0282739P.
PR 10-APR-2001; 2001US-0283112P.
PR 25-APR-2001; 2001US-0286630P.
XX (AGEN-) AGENSYS INC.
XX Jakobovits A, Challita-Eid PM, Paris M, Ge W, Hubert RS;
PI Morrison K, Morrison RK, Raitano AB;
XX WPI; 2003-075555/07.
DR N-PSDB; ABZ78140.
XX New composition comprising a substance that modulates the structure of
PT proteins and polynucleotides, useful for therapeutic, prognostic and
PT diagnostic reagents for eliciting cellular or humoral immune response in
PT cancer patients.
XX Claim 12; Fig 2R; 1021pp; English.
XX The present invention relates to novel human cancer-related genes and
CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
CC proteins are useful for eliciting a humoral or cellular immune response.
CC The genes are useful as probes and primers for the amplification and/or
CC detection of genes, mRNAs or their fragments, as reagents for the
CC diagnosis and/or prognosis of cancer, as coding sequences capable of
CC directing the expression of the protein, as tools for modulating or
CC inhibiting the expression of genes and/or translation of transcripts, and

CC as therapeutic agents. The proteins and peptides are useful as
CC therapeutic, prognostic and diagnostic reagents for cancer
XX
SQ Sequence 284 AA;

Alignment Scores:

Pred. No.: 3.01e-136 Length: 284
Score: 1530.00 Matches: 284
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 34.8% Indels: 0
DB: 6 Gaps: 0

US-10-768-158-1 (1-2419) x ABR01809 (1-284)

QY 21 ATGCGGAGAGCAGGCGGAGAGACCCAGCAGCCCGGGGAGTTCGAGACAGTACTTC 80
Db 1 MetAlaGluSerGluAlaGluThrProSerThrProGlyGluPheGluSerLysTrpPhe 20
QY 81 GAGTTCATGCGTGGCGCTGCCCGCTTCTGCCGCGGGAAGATGGAGGAGATCGCCCAAC 140
Db 21 GluPheHisGlyValArgLeuProPheCysArgGlyLysMetGluGluIleAlaAsn 40
QY 141 TTCCCGTGGCGGCCAGCGACGTGTGGATCGTCACCTACCCAGTCCCGCACCAGCTTG 200
Db 41 PheProValArgProSerAspValTrpIleValThrTyrProLysSerGlyThrSerLeu 60
QY 201 CTGACGAGGTGTCTACTTGGTGAGCCAGGGCGCTGACCCCGATGAGATCGGCTTGATG 260
Db 61 LeuGlnGluValTyrLeuValSerGlnGlyAlaAspProAspGluIleGlyLeuMet 80
QY 261 AACATCGACGAGCAGCTCCCGTCTGGAGTACCCACAGCCGGCGCTGGACATCATCAAG 320
Db 81 AsnIleAspGluGlnLeuProValLeuGlyTyrProGlnProGlyLeuAspIleIleLys 100
QY 321 GAATGACCTCTCCCGCTCATCAAGAGCCACCTGCCCTTACCGCTTTCTGCCCTCTGAC 380
Db 101 GluLeuThrSerProArgLeuIleLysSerHisLeuProTyrArgPheLeuProSerAsp 120
QY 381 CTCACATGAGACTCCAAAGTCTATATATGGCTCGGAACCCCAAGGATCTGGTGGTG 440
Db 121 LeuHisAsnGlyAspSerLysValIleTyrMetAlaArgAsnProLysAspLeuValVal 140
QY 441 TCTTATTATCAGTTCACACGCTCTCTGCGGACCATGAGTACCGAGGACCTTTCAAGAA 500
Db 141 SerTyrTyrGlnPheHisArgSerLeuArgThrMetSerTyrArgGlyThrPheGlnGlu 160
QY 501 TTCTGCCGAGGTTTATGAATGATAAGCTGGGCTACGGCTCCTGGTTTGGACAGCTCAG 560
Db 161 PheCysArgArgPheMetAsnAspLysLeuGlyTyrGlySerTrpPheGluHisValGln 180
QY 561 GAGTTCGGGAGCACCAGCATGGAGCTCGAAGCTGCTTTTCTCAAGTATGAAGACATGCAT 620
Db 181 GluPheTrpGluHisArgMetAspSerAsnValLeuPheLeuLysTrpGluAspMetHis 200
QY 621 CGGACCTGGTGGAGCTGGTGGAGCAGCTGCCAGATTCTCTGGGGGTGCTCTGTGACAAG 680
Db 201 ArgAspLeuValThrMetValGluGlnLeuAlaArgPheLeuGlyValSerCysAspLys 220
QY 681 GCCAGCTGGAAGCCCTGACGAGCAGCTGCCACAGCTGGTGGACCAAGTGTGCAACGCT 740
Db 221 AlaGlnLeuGluAlaLeuThrGluHisCysHisGlnLeuValAspGlnCysCysAsnAla 240
QY 741 GAGCCCTGCCGTCGGCGGGAGAGTGGGCTGTGGAGGACATCTTCACCGTCTCC 800
Db 241 GluAlaLeuProValGlyArgGlyArgValGlyLeuTrpLysAspIlePheThrValSer 260
QY 801 ATGAATGAGAAGTTGACTTGGTGTATATAACAGAGAAGATGGAAAGTGTGACCTCACGTTT 860
Db 261 MetAsnGluLysPheAspLeuValTyrLysGlnLysMetGlyLysCysAspLeuThrPhe 280
QY 861 GACTTTTATTTA 872
|||||

Db	281 AspPheTyrLeu 284	US-10-768-158-1 (1-2419) x ADQ89158 (1-284)
RESULT 4		
ADQ89158		
ID	ADQ89158 standard; protein; 284 AA.	
XX		
AC	ADQ89158;	
XX		
DT	21-OCT-2004 (first entry)	
XX		
DE	Human urological disorder related protein 16386 SEQ:110.	
XX		
KW	urological disorder; uropathic; cytostatic; urinary incontinence;	
KW	benign prostatic hyperplasia; human.	
XX		
OS	Homo sapiens.	
XX		
PN	WO2004065576-A2.	
XX		
PD	05-AUG-2004.	
XX		
PF	14-JAN-2004; 2004WO-US000750.	
XX		
PR	15-JAN-2003; 2003US-0440318P.	
PR	04-FEB-2003; 2003US-0444783P.	
PR	27-MAR-2003; 2003US-0457901P.	
PR	08-MAY-2003; 2003US-0468775P.	
PR	19-MAY-2003; 2003US-0471614P.	
PR	16-JUN-2003; 2003US-0478742P.	
PR	18-JUL-2003; 2003US-0488529P.	
PR	30-JUL-2003; 2003US-0491156P.	
PR	02-SEP-2003; 2003US-049594P.	
PR	26-SEP-2003; 2003US-0506332P.	
XX		
PA	(MILL-) MILLENNIUM PHARM INC.	
XX		
PI	Karicheti V, Silos-Santiago I, Eliasof SD;	
XX		
DR	N-PSDB; ADQ89157.	
XX		
PT	Use of polypeptides related to urological disorders, e.g. 44390, 54181,	
PT	211 or for identifying a compound capable of treating a urological	
PT	disorder or identifying and treating a subject having a urological	
XX	disorder.	
PS	Claim 1; SEQ ID NO 110; 542pp; English.	
XX		
CC	The present invention describes the use of polypeptides related to	
CC	urological disorders for identifying a compound capable of treating a	
CC	urological disorder, identifying a subject having a urological disorder,	
CC	or treating a subject having a urological disorder. Also described: (1) a	
CC	method for identifying a compound capable of treating a urological	
CC	disorder; (2) a method for identifying a subject having a urological	
CC	disorder; and (3) a method for treating a subject having a urological	
CC	disorder. The compound has uropathic and cytostatic activities. The	
CC	polypeptides related to urological disorders are useful for identifying a	
CC	compound capable of treating a urological disorder, identifying a subject	
CC	having a urological disorder, or treating a subject having a urological	
CC	disorder. Disorders include urinary incontinence and benign prostatic	
CC	hyperplasia. The present sequence represents a human urological disorder	
CC	related protein, which is used in the exemplification of the present	
XX	invention.	
XX		
SEQ	Sequence 284 AA;	
Alignment Scores:		
Pred. No.:	3,01e-136	Length: 284
Score:	1530.00	Matches: 284
Percent Similarity:	100.0%	Conservative: 0
Best Local Similarity:	100.0%	Mismatches: 0
Query Match:	34.8%	Indels: 0
DB:	8	Gaps: 0

RESULT 5
ADR44884
ID ADR44884 standard; protein; 284 AA.
XX
AC ADR44884;
XX
DT 18-NOV-2004 (first entry)
XX
XX Polypeptide 16386 amino acid sequence.
XX

analgesic; gene therapy; pain; painful disorder.

XX Homo sapiens.

XX WO2004071411-A2.

XX 26-AUG-2004.

XX 30-JAN-2004; 2004WO-US002851.

XX 04-FEB-2003; 2003US-0444781P.

XX 05-MAR-2003; 2003US-0452291P.

XX 13-MAR-2003; 2003US-0454540P.

XX 16-JUN-2003; 2003US-0478805P.

XX 30-JUL-2003; 2003US-0491048P.

XX (MILL-) MILLENNIUM PHARM INC.

XX Silos-Santiago I, Karicheti V, Eliasof SD;

XX WPI; 2004-625760/60.

XX N-PSDB; ADR44883.

XX Identifying a compound capable of treating pain by combining a compound to be tested with e.g., 16386 polypeptide or with a host cell expressing the polypeptide and detecting the binding of the test compound to the polypeptide.

XX Disclosure; SEQ ID NO 2; 233pp; English.

XX The invention relates to a method of identifying a compound capable of treating pain or painful disorder by combining a compound to be tested with an e.g., 16386, 15402, 21165, 1423, 636, or 32620 polypeptide or with a host cell expressing the polypeptide and detecting the binding of the test compound to the polypeptide to identify a compound that binds to the polypeptide. The method is useful in identifying a compound capable of treating pain or painful disorder. This sequence corresponds to the polypeptide 16386 - a brain sulfotransferase-like protein.

XX Alignment Scores:

XX Pred. No.: 3.01e-136 Length: 284

XX Score: 1530.00 Matches: 284

XX Percent Similarity: 100.0% Conservative: 0

XX Best Local Similarity: 100.0% Mismatches: 0

XX Query Match: 34.8% Indels: 0

XX DB: 8 Gaps: 0

XX US-10-768-158-1 (1-2419) x ADR44884 (1-284)

QY 21 ATGCGGAGAGCGAGCGGAGACCCAGCACCCCGGGGAGTTCGAGAGCAAGTACTTC 80

Db 1 MetAlaGluSerGluAlaGluThrProSerThrProGlyGluPheGluSerLysTyrPhe 20

QY 81 GAGTTCCATGGGTGGCGCTCCGCGCTTCGCGCGGGAAGATGGAGAGATCGCCAAC 140

Db 21 GluPheHisGlyValArgLeuProProPheCysArgGlyLysMetGluGluAlaAsn 40

QY 141 TTCCCGGTGCGGCGGAGCGAGTGTGGATCGTCACCTACCCCAAGTCGCGGACCAAGCTTG 200

Db 41 PheProValArgProSerAspValTrpIleValThrTyrProLysSerGlyThrSerLeu 60

QY 201 CTGACGAGGTGGTCTACTTGGTAGCCAGGCGCTGACCCCGATGATCGGCTTGATG 260

Db 61 LeuGlnGluValValTyrLeuValSerGlnGlyAlaAspProAspGluIleGlyLeuMet 80

QY 261 AACATCGACGAGCTCCCGTCTGGAGTACCCACAGCCGGCGCTGCACATCATCAAG 320

Db 81 AsnIleAspGluGlnLeuProValLeuGluTyrProGlnProGlyLeuAspIleLeuLys 100

QY 321 GAATGACCTCTCCCGGCTCATCAAGAGCCACCTGACCCGCTTCTGCCCTCTGAC 380

Db 101 GluLeuThrSerProArgLeuIleLysSerHisLeuProTyrArgPheLeuProSerAsp 120

QY 381 CTCACAATGGAGACTCCAAGTCACTATATGGTCGCAACCCCAAGGATCTCTGGTGGTG 440

Db 121 LeuHisAsnGlyAspSerLysValIleTyrMetAlaArgAsnProLysAspLeuValVal 140

QY 441 TCTTATTATCAGTTCCACCGCTCTCTCGGACCATGAGCTACCGAGGACACCTTTCAAGAA 500

Db 141 SerTyrTyrGlnPheHisArgSerLeuArgThrMetSerTyrArgGlyThrPheGlnGlu 160

QY 501 TTCTGCCGAGGTTTATGAATGATAGCTGGCTACCGCTCTCTGGTTTGACACGTCGAG 560

Db 161 PheCysArgArgPheMetAsnAspLysLeuGlyTyrGlySerTyrPheGluHisValGln 180

QY 561 GAGTTCTGGAGCACCGCATGGACTCGAAGCTGCTTTTCTCAAGTATGAAGACATGCAT 620

Db 181 GluPheTrpGluHisArgMetAspSerAsnValLeuPheLeuLysTyrGluAspMetHis 200

QY 621 CGGACCTGGTGCATGGTGGAGCAGCTGCCAGATTCTCTGGGGGTGTCTGTGCAAG 680

Db 201 ArgAspLeuValThrMetValGluGlnLeuAlaArgPheLeuGlyValSerCysAspLys 220

QY 681 GCCAGCTGGAAGCCCTGACGGAGCAGCTGCACAGCTGGTGGACCAAGTCTGCAACGCT 740

Db 221 AlaGlnLeuGluAlaLeuThrGluHisCysHisGlnLeuValAspGlnCysCysAsnAla 240

QY 741 GAGCCCTGCGCGTGGCGGGGAGAGTTGGGCTGTGGAGGACATCTTCACCGTCTCC 800

Db 241 GluAlaLeuProValGlyArgGlyArgValGlyLeuTrpLysAspIlePheThrValSer 260

QY 801 ATGAATCAGAAAGTTTGACTTGGTGTATAAACAAGAGATGGAAAGTGTGACCTCACGTTT 860

Db 261 MetAsnGluLysPheAspLeuValTyrLysGlnLysMetGlyLysCysAspLeuThrPhe 280

QY 861 GACTTTTATTATTA 872

Db 281 AspPheTyrLeu 284

RESULT 6

ADZ21933

ID ADZ21933 standard; protein; 284 AA.

XX AC ADZ21933;

XX DT 16-JUN-2005 (first entry)

XX DE Human SULT4A1 splice variant 1 SEQ ID NO 1.

XX KW neuroprotective; nootropic; diagnosis; prognosis; gene expression;

XX KW neurodegenerative disease; neurological disease; Alzheimers disease;

XX KW degeneration; sulfotransferase family 4A member 1; SULT4A1;

XX KW splice variant; enzyme.

XX OS Homo sapiens.

XX PN WO2005030947-A2.

XX PD 07-APR-2005.

XX PF 29-SEP-2004; 2004WO-EP052353.

XX PR 30-SEP-2003; 2003US-0506775P.

XX PA (EVOT-) EVOTEC NEUROSCIENCES GMBH.

XX PI Von Der Kammer H, Pohlner J;

XX WPI; 2005-262951/27.

XX N-PSDB; ADZ21938, ADZ21935.

XX Diagnosing or prognosticating a neurodegenerative disease in a subject by determining a level and/or activity of a transcription and/or translation product of a gene coding for a cytosolic sulfotransferase family 4A

PT member 1.
 XX Claim 15; SEQ ID NO 1; 83pp; English.
 XX

The invention describes a method of diagnosing or prognosticating a neurodegenerative disease in a subject, or determining whether a subject is at increased risk of developing the disease. The method comprises determining a level and/or an activity of a transcription product and/or a translation product of a gene coding for a cytosolic sulfotransferase family 4A member 1 and/or a fragment, or derivative or variant of the transcription or translation product, in a sample obtained from the subject and comparing the level and/or the activity to a reference value representing a known disease or health status, therefore diagnosing or prognosticating the neurodegenerative disease in the subject, or determining whether the subject is at increased risk of developing the neurodegenerative disease. Protein molecules comprising fully defined 284 given in the specification, the protein molecules being translation products of the gene coding for a cytosolic sulfotransferase family 4A member 1, or its fragments, or derivatives, or variants, are useful as diagnostic targets for detecting a neurodegenerative disease, preferably Alzheimer's disease or as screening targets for reagents or compounds preventing, or treating, or ameliorating the disease. An antibody specifically immunoreactive with an immunogen, where the immunogen is a translation product of a gene coding for a cytosolic sulfotransferase family 4A member 1, SEQ ID NO. 1 or 2, or its fragment, or derivative, or variant is useful for detecting the pathological state of a cell in a sample obtained from a subject, comprising immunocytochemical staining of the cell with the antibody, where an altered degree of staining, or an altered staining pattern in the cell compared to a cell representing a known health status indicates a pathological state of the cell which relates to a neurodegenerative disease, preferably to Alzheimer's disease. The genetically altered non-human animal is useful for screening, testing, and validating compounds, agents, and modulators in the development of diagnostics and therapeutics to treat neurodegenerative diseases, in particular Alzheimer's disease. An agent(s) which directly or indirectly affect an activity and/or a level of a gene coding for a cytosolic sulfotransferase family 4A member 1, and/or a transcription product or a translation product of a gene coding for a cytosolic sulfotransferase family 4A member 1, and/or their fragment, or derivative, or variant is useful for treating or preventing a neurodegenerative disease, in particular Alzheimer's disease. This is the amino acid sequence of human sulfotransferase family 4A member 1 splice variant 1 (SULT4A1sv1).

XX
 SQ Sequence 284 AA;

Alignment Scores:
 Pred. No.: 3.01e-136 Length: 284
 Score: 1530.00 Matches: 284
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 34.8% Indels: 0
 DB: 9 Gaps: 0

US-10-768-158-1 (1-2419) x ADZ21933 (1-284)

Qy 21 ATGCGGAGAGCGAGCGGAGACCCCGAGACCCCGGGGAGTTTCGAGAGCAAGTACTTC 80
 Db 1 MetAlaGluSerGluAlaGluThrProSerThrProGlyGluPheGluSerLysTyPhe 20
 Qy 81 GAGTTCCATGCGGTGCGGCTGCCGCCCTTCGCGCGGGAGAGATGGAGGATCGCCAAAC 140
 Db 21 GluPheHisGlyValArgLeuProProPheCysArgGlyLysMetGluGluLeuAlaAsn 40
 Qy 141 TTCCCGGTGGCGGCCAGGAGCGTGTGATCTGATACCTACCCCAAGTCGGGACGAGTTG 200
 Db 41 PheProValArgProSerAspValTrpIleValThrTyProLysSerGlyThrSerLeu 60
 Qy 201 CTGCAGAGGTGGTCTACTTGGTCAGCCAGGCGCTGACCCCGATGATCGGCTTCATG 260
 Db 61 LeuGlnGluValValTyLeuValSerGlnGlyAlaAspProAspGluIleGlyLeuMet 80

Qy 261 AACATCAGCAGCAGCAGCTCCCGGTCTCGAGTACCCACAGCCGGGCTTGACATCATCAAG 320
 Db 81 AsnIleAspGluGlnLeuProValLeuGluTyProGlnProGlyLeuAspIleLeuLys 100
 Qy 321 GAACTGACCTCTCCCGCCTCATCAAGAGCCACCTGCCCTACCGCTTTCTGCCCTCTGAC 380
 Db 101 GluLeuThrSerProArgLeuIleLysSerHisLeuProTyArgPheLeuProSerAsp 120
 Qy 381 CTCACAAATGAGACTCCAAAGGTCACTATATGCTCGCAACCCCAAGGATCTCGTGGTG 440
 Db 121 LeuHisAsnGlyAspSerLysValIleTyMetAlaArgAsnProLysAspLeuValVal 140
 Qy 441 TCTATTATCAGTTCCACCCCTCTCGCGACCATAGCTACCGAGGACACCTTTTCAGAA 500
 Db 141 SerTyTyArgGlnPheHisArgSerLeuArgThrMetSerTyArgGlyThrPheGlnGlu 160
 Qy 501 TTCTGCCGAGGTTTATGAATGATGAAGCTGGGCTACGGCTCCTGGTTTGGACAGCTGCAG 560
 Db 161 PheCysArgArgPheMetAsnAspLysLeuGlyTyArgSerTrpPheGluHisValGln 180
 Qy 561 GAGTTCTGGGAGCACCACATGGACTCGAACGTGCTTTTCTCAAGTATGAAGACATGCAT 620
 Db 181 GluPheTrpGluHisArgMetAspSerAsnValLeuPheLeuLysTyArgGlyMetHis 200
 Qy 621 CGGACCTGTGAGCAGTGGTGGAGCAGCTGCCAGATTCCTGGGGTGTCTCTGTGACAAG 680
 Db 201 ArgAspLeuValThrMetValGluGlnLeuAlaArgPheLeuGlyValSerCysAspLys 220
 Qy 681 GCCAGCTGGAAGCCCTGACGGAGCACTGCCACAGCTGGTGGACCAAGTGTCTCAACGCT 740
 Db 221 AlGlnLeuGluAlaLeuThrGluHisCysHisGlnLeuValAspGlnCysCysAsnAla 240
 Qy 741 GAGGCCCTGCGCGTGGGCGCGGGAAGATTGGGCTGTGGAAGGACATCTTCACCGTCTCC 800
 Db 241 GluAlaLeuProValGlyArgGlyArgValGlyLeuTrpLysAspIlePheThrValSer 260
 Qy 801 ATGAATGAGAAGTTGACTGGTGTATAAACAGAGATGGAAAGTGTGACCTCAGCTTT 860
 Db 261 MetAsnGluLysPheAspLeuValTyLysGlnLysMetGlyLysCysAspLeuThrPhe 280
 Qy 861 GACTTTTATTATA 872
 Db 281 AspPheTyLeu 284

RESULT 7
 ABM04807
 ID ABM04807 standard; protein; 284 AA.
 XX AC ABM04807;
 XX DT 22-SEP-2003 (first entry)
 XX DE Rat sulfotransferase-like protein.
 XX KW spinal cord; neuropathic pain; central sensitisation pain; pain;
 XX KW analgesic; gene therapy.
 XX OS Rattus norvegicus.
 XX PN EP1284298-A2.
 XX PD 19-FEB-2003.
 XX PF 26-JUL-2002; 2002EP-00255229.
 XX PR 27-JUL-2001; 2001GB-00018354.
 XX PR 07-FEB-2002; 2002GB-00002883.
 XX PA (WARN) WARNER LAMBERT CO.
 XX PI Brooksbank RA, Dixon AK, Lee K, Pinnock RD;
 XX WPI; 2003-543489/52.
 DR

Query Match:	26.0%	Indels:	45
DB:	6	Gaps:	1
US-10-768-158-1 (1-2419) x ABR01861 (1-218)			
QY	88	ATGGCGTGGCGTGGCGCCCTTCTGCGCGGGAGATGGAGAGATCGCCAACTTCCCGG	147
Db	1	MetAlaCysGlyCysArgProSerAlaAlaGlyArgArgSerProThrSerArg	20
QY	148	TGCGGCCCGAGCGCTGTGATCGTCACTACCCCAAGTCGGCACCAGCTTGCTGCAGG	207
Db	21	CysGlyProAlaThrCysGlySerSerProThrProSerPro-----	34
QY	208	AGTGGTCTACTTGGTGAGCCAGGCGCTGACCCCGATGAGATCGGCTTGATGAACATCG	267
Db	34	-----	34
QY	268	ACGAGCAGCTCCCGTCTCGAGTACCCACAGCGGGCTGGACATCATCAAGAACTGA	327
Db	35	-----Glu--LeuT	37
QY	328	CCTCCTCCCGCTCATCAAGACCACCTGCCTTCTGCGCTTCTGACCTCACA	387
Db	37	hrSerProArgLeuIleIleSerHisLeuProTyrArgPheLeuProSerAspLeuHisA	57
QY	388	ATGGAGACTCCAAGGTCATCTATATGGCTCGCAACCCCAAGGATCTGGTGTGCTTATT	447
Db	57	enGlyAspSerLeuValIleTyrMetAlaArgAsnProLysAspLeuValSerTyrT	77
QY	448	ATCATGTTCCCGCTCTCGGGACCATGAGTACCCAGGACCTTTCAAGATTCGCC	507
Db	77	yrGlnPheHisArgSerLeuArgThrMetSerTyrArgGlyThrPheGlnGluPheCysA	97
QY	508	GGAGTTTATGAATGATAAGCTGGGCTACGCTCTGTTTGGACAGCTGCAGGAGTTCT	567
Db	97	rgArgPheMetAsnAspLysLeuGlyTyrGlySerTrpPheGluHisValGlnGluPheT	117
QY	568	GGAGCACCCGATGGATCGAACTGCTTTTCTCAAAGTATGAAGACATCATCGGAC	627
Db	117	rpGluHisArgMetAspSerAsnValLeuPheLeuLysTyrGluAspMetHisArgAspL	137
QY	628	TGTTGAGATGGTGGACAGCTGCCAGATCTTGGGGGTGCTCTGTGACAAGCCCGAC	687
Db	137	euValThrMetValIleGlnLeuAlaArgPheLeuGlyValSerCysAspLysAlaGlnL	157
QY	688	TGGAAGCCCTGACGAGCAGCTGCCACAGCTGGTGACCACTGTCTGCAACGCTGAGGCC	747
Db	157	euGluAlaLeuThrGluHisCysHisGlnLeuValAspGlnCysAsnAlaGluAlaL	177
QY	748	TGCGCGTGGCGGGGAAGATTTGGGCTGTGGAAGGACATCTTCCCGTCTCCATGAATG	807
Db	177	euProValGlyArgGlyArgValGlyLeuTrpLysAspIlePheThrValSerMetAsnG	197
QY	808	AGAAGTTTGACTTGCTGTATAACAGAAAGATGGGAAAGTGTGACCTCAGCTTGACTTTT	867
Db	197	lulysPheAspLeuValTyrLysGlnLysMetGlyLysCysAspLeuThrPheAspPheT	217
QY	868	ATTTA 872	
Db	217	yrLeu 218	
RESULT 9			
ABR01861			
ID	ABR01861 standard; protein; 171 AA.		
XX	AC ABR01861;		
XX	19-MAY-2003 (first entry)		
XX	Human cancer-related protein, 192p2G7 v.3.		
XX	Human; cytostatic; vaccine; cancer; immune response.		
XX			

OS	Homo sapiens.		
XX	WO200283921-A2.		
XX	24-OCT-2002.		
PD			
XX	10-APR-2002; 2002WO-US011654.		
PF			
XX	10-APR-2001; 2001US-0282739P.		
PR	10-APR-2001; 2001US-0283112P.		
PR	25-APR-2001; 2001US-0286630P.		
XX	(AGEN-) AGENSYS INC.		
PA			
XX	Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;		
PI	Morrison K, Morrison RK, Raitano AB;		
PI			
XX	WPI; 2003-075555/07.		
DR	N-PSDB; ABZ78168.		
DR			
XX	New composition comprising a substance that modulates the structure of		
PT	proteins and polynucleotides, useful for therapeutic, prognostic and		
PT	diagnostic reagents for eliciting cellular or humoral immune response in		
PT	cancer patients.		
XX			
XX	Example 5; Fig 12r; 102lpp; English.		
PS			
XX	The present invention relates to novel human cancer-related genes and		
CC	proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and		
CC	proteins are useful for eliciting a humoral or cellular immune response.		
CC	The genes are useful as probes and primers for the amplification and/or		
CC	detection of genes, mRNAs or their fragments, as reagents for the		
CC	diagnosis and/or prognosis of cancer, as coding sequences capable of		
CC	directing the expression of the protein, as tools for modulating or		
CC	inhibiting the expression of genes and/or translation of transcripts, and		
CC	as therapeutic agents. The proteins and peptides are useful as		
CC	therapeutic, prognostic and diagnostic reagents for cancer		
XX			
SQ	Sequence 171 AA;		
Alignment Scores:			
Pred. No.:	3,48e-73	Length:	171
Score:	867.50	Matches:	170
Percent Similarity:	60.2%	Conservative:	1
Best Local Similarity:	59.9%	Mismatches:	0
Query Match:	19.7%	Indels:	113
DB:	6	Gaps:	1
US-10-768-158-1 (1-2419) x ABR01861 (1-171)			
QY	21	ATGCGGAGAGCGAGCGCGGAGACCCAGACACCCCGGGGAGTTTCGAGAGCAAGTACTTC	80
Db	1	MetAlaGluSerGluAlaGluThrProSerThrProGlyGluPheGluSerLysTyrPhe	20
QY	81	GAGTTCATGCGCTGGCGGTGCGCCCTTCTGCGCGGGAAGATGGAGGAGATCGCCAAC	140
Db	21	GluPheHisGlyValArgLeuProPheCysArgGlyLysMetGluGluIleAlaAsn	40
QY	141	TTCCCGTGGCGGCCAGCGAGCTGTGATCGTCACTACCTACCCCAAGTCCGGCACCAGCTTG	200
Db	41	PheProValArgProSerAspValTrpIleValThrTyrProLysSer-----	56
QY	201	CTGACGAGGTGTCTACTTGGTGAGCCAGGGCGCTGACCCCGATGAGATCGGCTTGATG	260
Db	56	-----	56
QY	261	AACATGACGAGCAGCTCCCGTCTCGGAGTACCCACAGCGGGCGCTGGACATCATCAAG	320
Db	56	-----	56
QY	321	GAAGTACCTCTCCCGCTCATCAAGAGCCACCTGCCCTTACCGCTTCTGCGCCTCTGAC	380
Db	56	-----	56

QY 381 CTCCACATGGAGCTCCAGGTCTATATGCTCGCAACCCCAAGGATCTGGTGGT 440
 Db 56 ----- 56
 QY 441 TCTATTATCAGTTCACCGCTCTCTCGGACCATGAGCTACCGAGGCACCTTCAAGAA 500
 Db 56 ----- 56
 QY 501 TTCTGCGGAGGTTTATGAATGATAAGCTGGCTACCGCTCCTGGTTTGAGCAGTGCAG 560
 Db 57 -----ValGlyTyrGlySerTrpPheGluHisValGln 67
 QY 561 GAGTCTTGGGAGCACCAGCTCGACTCGAAGCTGCTTTTCTCAAGTATGAAGATCCAT 620
 Db 68 GluPheTrpGluHisArgMetAspSerAsnValLeuPheLeuLysTrpGluAspMetHis 87
 QY 621 CGGACCTGGTGACGATGGTGGAGCAGCTGCCAGATTCCTGGGGGTGCTCTGTGACAAG 680
 Db 88 ArgAspLeuValThrMetValGluGlnLeuAlaArgPheLeuGlyValSerCysAspLys 107
 QY 681 GCCAGCTGAAGCCCTGACGGAGCATGTCACAGCTGGTGACCAAGTGTGCAAGCT 740
 Db 108 AlaGlnLeuGluAlaLeuThrGluHisCysHisGlnLeuValAspGlnCysAsnAla 127
 QY 741 GAGGCCCTGCGCTGGCGCGGAGAGTGGCTGTGGAAGGACATCTTCACGGTCTCC 800
 Db 128 GluAlaLeuProValGlyArgValGlyLeuTrpLysAspPheThrValSer 147
 QY 801 ATGAATCAGAAGTTTGACTTGGTGTATAAACAGAAAGATGGAAAGTGTGACCTCAGCTTT 860
 Db 148 MetAsnGlnLysPheAspLeuValTyrLysGlnLysMetGlyLysCysAspLeuThrPhe 167
 QY 861 GACTTTTATTTA 872
 Db 168 AspPheTyrLeu 171
 RESULT 10
 ADZ21934
 ID ADZ21934 standard; protein; 171 AA.
 XX AC ADZ21934;
 DX 16-JUN-2005 (first entry)
 XX Human SULT4A1 splice variant 2 SEQ ID NO 2.
 KW neuroprotective; nootropic; diagnosis; prognosis; gene expression;
 KW neurodegenerative disease; neurological disease; Alzheimers disease;
 KW degeneration; sulfotransferase family 4A member 1; SULT4A1;
 KW splice variant; enzyme.
 XX OS Homo sapiens.
 XX WO2005030947-A2.
 XX PN 07-APR-2005.
 XX PD 29-SEP-2004; 2004WO-EP052353.
 XX PF 30-SEP-2003; 2003US-0506775P.
 XX PR (EVOT-) EVOTEC NEUROSCIENCES GMBH.
 XX PA Von Der Kammer H, Pohlner J;
 XX PI WPI; 2005-262951/27.
 XX DR N-PSDB; ADZ21936.
 XX PT Diagnosing or prognosticating a neurodegenerative disease in a subject by
 PT determining a level and/or activity of a transcription and/or translation
 PT product of a gene coding for a cytosolic sulfotransferase family 4A
 PT member 1.

XX Claim 15; SEQ ID NO 2; 83pp; English.
 PS The invention describes a method of diagnosing or prognosticating a
 XX neurodegenerative disease in a subject, or determining whether a subject
 CC is at increased risk of developing the disease. The method comprises
 CC determining a level and/or an activity of a transcription product and/or
 CC a translation product of a gene coding for a cytosolic sulfotransferase
 CC family 4A member 1 and/or a fragment, or derivative or variant of the
 CC subject and comparing the level and/or the activity to a reference value
 CC representing a known disease or health status, therefore diagnosing or
 CC prognosticating the neurodegenerative disease in the subject, or
 CC determining whether the subject is at increased risk of developing the
 CC neurodegenerative disease. Protein molecules comprising fully defined 284
 CC and/or 171 amino acid (SEQ ID NO. 1 and/or 2, respectively) sequence
 CC given in the specification, the protein molecules being translation
 CC products of the gene coding for a cytosolic sulfotransferase family 4A
 CC member 1, or its fragments, or derivatives, or variants, are useful as
 CC diagnostic targets for detecting a neurodegenerative disease, preferably
 CC Alzheimer's disease or as screening targets for reagents or compounds
 CC preventing, or treating, or ameliorating the disease. An antibody
 CC specifically immunoreactive with an immunogen, where the immunogen is a
 CC translation product of a gene coding for a cytosolic sulfotransferase
 CC family 4A member 1, SEQ ID NO. 1 or 2, or its fragment, or derivative, or
 CC variant is useful for detecting the pathological state of a cell in a
 CC sample obtained from a subject, comprising immunocytochemical staining of
 CC the cell with the antibody, where an altered degree of staining, or an
 CC altered staining pattern in the cell compared to a cell representing a
 CC known health status indicates a pathological state of the cell which
 CC relates to a neurodegenerative disease, preferably to Alzheimer's
 CC disease. The genetically altered non-human animal is useful for
 CC screening, testing, and validating compounds, agents, and modulators in
 CC the development of diagnostics and therapeutics to treat
 CC neurodegenerative diseases, in particular Alzheimer's disease. An
 CC agent(s) which directly or indirectly affect an activity and/or a level
 CC of a gene coding for a cytosolic sulfotransferase family 4A member 1,
 CC and/or a transcription product or a translation product of a gene coding
 CC for a cytosolic sulfotransferase family 4A member 1, and/or their
 CC fragment, or derivative, or variant is useful for treating or preventing
 CC a neurodegenerative disease, in particular Alzheimer's disease. This is
 CC the amino acid sequence of human sulfotransferase family 4A member 1
 CC splice variant 2 (SULT4A1sav1).
 XX
 SQ Sequence 171 AA;
 Alignment Scores:
 Pred. No.: 3,48e-73 Length: 171
 Score: 867.50 Matches: 170
 Percent Similarity: 60.2% Conservative: 1
 Best Local Similarity: 59.9% Mismatches: 0
 Query Match: 19.7% Indels: 113
 DB: Gaps: 1
 US-10-768-158-1 (1-2419) x ADZ21934 (1-171)
 QY 21 ATGCGGAGAGCGAGCGCGAGACCCCGAGACCCCGGGGAGTTCGAGACAGTACTTC 80
 Db 1 MetAlaGluSerGluAlaGluThrProSerThrProGlyGluPheGluSerLysTyrPhe 20
 QY 81 GAGTTCCATGGCTGGCGCTGCCGCCCTTCTGCGGGGGAAGATGGAGGAGATGCCAAC 140
 Db 21 GluPheHisGlyValArgLeuProProPheCysArgGlyLysMetGluGluIleAlaAsn 40
 QY 141 TTCCCGGTGCGCCCGAGCGACGTGTGGATCGTCACTTACCCCAAGTCCCGACAGCTTG 200
 Db 41 PheProValArgProSerAspValTyrPheValThrTyrProLysSer----- 56
 QY 201 CTCGAGGAGTGTGCTCTACTTGTGTGAGCCCGGCGCTGACCCCGATGAGATCGGCTTGATG 260
 Db 56 ----- 56
 QY 261 AACATCGACGAGCAGCTCCCGGTCTGTGAGTACCCACAGCCGGCGCTTGGACATCATCAAG 320

Db 56 ----- 56
Qy 321 GAAGTACCTCTCCCGCCTCATCAAGAGCACCTGCCTTCTGCGCCTCTGAC 380
Db 56 ----- 56
Qy 381 CTCACAATGGAGACTCCAAGGTCATCTATATGCTCGCAACCCCAAGGATCTGGTGTG 440
Db 56 ----- 56
Qy 441 TCTTATTATCAGTTCCACCGCTCTCTCGGACCATGAGTACCGAGGCACCTTTCAGAA 500
Db 56 ----- 56
Qy 501 TTCTGCGGAGGTTTATGAATGATAAGCTGGGTACCGGCTCTGTTTGGACAGTGCAG 560
Db 57 -----ValGlyTyrGlySerTrpPheGluHisValGln 67
Qy 561 GAGTTCGGGAGCACCCGATCGAGCTCGAAGCTGCTTTTCTCAAGTATGAAGCATGCAT 620
Db 68 GluPheTrpGluHisArgMetAspSerAsnValLeuPheLeuLysTyrGluAspMetHis 87
Qy 621 CGGAGCTGTGAGTGTGGAGCAGCTGCCAGATTCCTGGGGTCTCTGTGACAAG 680
Db 88 ArgAspLeuValThrMetValGluGlnLeuAlaArgPheLeuGlyValSerCysAspLys 107
Qy 681 GCCAGCTGGAAGCCCTGACGGAGCAGCTGCCACAGCTGGTGACAGTGTGCAACGCT 740
Db 108 AlaGlnLeuGluAlaLeuThrGluHisCysHisGlnLeuValAspGlnCysCysAsnAla 127
Qy 741 GAGGCCCTGCCCTGGCCCGGGAAGAGTTGGGCTGTGGAAGGACAFCTTCACCGTCTCC 800
Db 128 GluAlaLeuProValGlyArgGlyArgValGlyLeuTrpLysAspIlePheThrValSer 147
Qy 801 ATCAATGAGAGCTTGACTTGGTGTATAAACAGAGATGGGAAGTGTGACCTCACGTTT 860
Db 148 MetAsnGluLysPheAspLeuValTyrLysGlnLysMetGlyLysCysAspLeuThrPhe 167
Qy 861 GACTTTTATTATA 872
Db 168 AspPheTyrLeu 171

RESULT 11

AAU78269
ID AAU78269 standard; protein; 159 AA.
AC AAU78269;
XX
DT 05-JUN-2002 (first entry)
XX
DE Rat nervous system-specific sulfotransferase, SULTn, protein sequence.
XX
KW Rat; nervous system-specific sulfotransferase; SULTn; sulfonation;
XX neurological disorder; multiple sclerosis; Huntington's chorea;
XX nervous system-mediated muscular atrophy; amyotrophic lateral sclerosis;
XX myasthenia gravis; Alzheimer's disease; senile dementia; Schizophrenia;
XX bipolar disorder; cytosolic sulfotransferase gene family;
XX cerebellar ataxia.
XX
OS Rattus sp.
XX
PN WO200218541-A2.
XX
XX
PD 07-MAR-2002.
XX
PF 04-SEP-2001; 2001WO-US041986.
XX
PR 01-SEP-2000; 2000US-0229929P.
XX
PA (UYBO-) UNIV BOSTON.
XX
XX Parb DH, Martin S;
PI

XX
DR WPI; 2002-281063/32.
XX P-PSDB; AAU78269.
PT New nervous system-specific sulfotransferase (SULTn) polypeptide for
PT treating neurological disorders including multiple sclerosis, nervous
PT system-mediated muscular atrophy and Huntington's chorea.
XX
PS Claim 2; Page; 46pp; English.
XX
CC The present invention relates to a new nervous system-specific
CC sulfotransferase (SULTn) polypeptide. The SULTn polypeptide of the
CC invention comprises a nervous system-specific sulfotransferase,
CC comprising a sequence of 284 amino acids, given in the specification. The
CC invention is useful for identifying an effector of nervous system-
CC specific sulfonation by contacting the polypeptide with a candidate
CC effector compound and assaying the ability of the candidate effector
CC compound to effect the sulfotransferase function of the polypeptide. The
CC invention is also useful for inhibiting nervous system-specific
CC sulfonation by contacting the polypeptide under physiological conditions
CC with a compound capable of interacting with and inhibiting the
CC sulfonation activity of the polypeptide, where the compound is an
CC antibody, peptide, polypeptide, nucleic acid, organic molecule or an
CC inorganic molecule. The polypeptide of the invention is also useful for
CC treating a neurological disorder, including multiple sclerosis, nervous
CC system-mediated muscular atrophy, amyotrophic lateral sclerosis,
CC Huntington's chorea, myasthenia gravis, Alzheimer's disease, senile
CC dementia, Schizophrenia, bipolar disorder and cerebellar ataxia. The
CC present amino acid sequence represents the rat SULTn protein of the
CC invention. SULTn is a member of the cytosolic sulfotransferase gene
CC family. Note: The sequence data for this sequence did not form part of
CC the printed specification but was obtained in electronic format from
CC GENBANK AF176343
XX
SQ Sequence 159 AA;

Alignment Scores:
Pred. No.: 1.4e-71 Length: 159
Score: 850.50 Matches: 157
Percent Similarity: 99.4% Conservative: 2
Best Local Similarity: 98.1% Mismatches: 0
Query Match: 19.3% Indels: 1
DB: 5 Gaps: 1

US-10-768-158-1 (1-2419) x AAU78269 (1-159)
Qy 393 GACTCCCAAGTCTATATGGCTCGCAACCCCAAGGATCTGGTGTCTTATTATCAG 452
Db 1 AspSerLysValIleTyrMetAlaArgAsnProLysAspLeuValSerTyrTyrGln 20
Qy 453 TTCACCGCTCTCTGCGGACCATGAGCTACCGAGGCACCTTTCAAGAATTTCTGCGGAGG 512
Db 21 PheHisArgSerLeuArgThrMetSerTyrArgGlyThrPheGlnGluPheCysArgArg 40
Qy 513 TTTATGAATGATTAAGCTGGGCTACCGCTCTCTGGTTTGTAGCACGTGACGAGTCTGGGAG 572
Db 41 PheMetAsnAspLysLeuGlyTyrGlySerTrpPheGluHisValGlnGluPheTrpGlu 60
Qy 573 CACCGCATGGACTCGAAGCGTCTTTTCTCAAGTATGACACATGCATCGGAGCCTCGTG 632
Db 61 HisArgMetAspAlaAsnValLeuPheLeuLysTyrGluAspMetHisArgAspLeuVal 80
Qy 633 ACATGCTGGAGCAGCTGGCCAGATTCTCTGGGGGTGTCTCTGTGACAAAGGCCAGCTGGAA 692
Db 81 ThrMetValGluGlnLeuAlaArgPheLeuGlyValSerCysAspLysAlaGlnLeuGlu 100
Qy 693 GCCCTGACGAGCAGCTGCCACCGCTGGTGGACAGTGTGTGCAACGCTGAGGCCCTGCC 752
Db 101 SerLeu---GluHisCysHisGlnLeuValAspGlnCysCysAsnAlaGluAlaLeuPro 119
Qy 753 GTGGGCGGGGAGAGTGTGGCTGTGGAAGGACATCTTCACCGTCTCCATGATGAGAG 812
Db 120 ValGlyArgGlyArgValGlyLeuTrpLysAspIlePheThrValSerMetAsnGluLys 139

PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249224P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0251990P.
PR 05-JAN-2001; 2001US-0259678P.
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX
XX
PI Rosen CA, Barash SC, Ruben SM;
DR WPI; 2001-488787/53.
XX N-PSDB; AAS30195.
XX
PT New polynucleotides and polypeptides, useful for diagnosing, treating,
PT preventing or prognosing e.g. kidney, cardiovascular, blood, electrolyte
PT imbalance or neoplastic disorders, autoimmune diseases, cancers.
XX
XX
PS Claim 1; SEQ ID NO 113; 506pp; English.
XX
XX
CC The invention relates to novel nucleic acids and polypeptides useful for
CC diagnosing, treating, preventing and/or prognosing disorders related to
CC these polypeptides. The polynucleotides are especially useful in the
CC diagnosis, prognosis, prevention and/or treatment of diseases which
CC include kidney disorders (e.g. renal failure or nephritis),
CC cardiovascular disorders (e.g. hypertension or myocardial infarction),
CC blood disorders (e.g. anaemia or blood coagulation disorders),
CC electrolyte imbalance disorders (e.g. hyponatraemia or hyperkalaemia),
CC neoplastic disorders (e.g. nephroma or renal cell cancer), autoimmune
CC diseases, cancers, inflammatory diseases, reproductive system disorders,
CC endocrine disorders, neural activity and neurological disorders, wound
CC healing and respiratory disorders. AAU18644-AAU18715 represent the novel
CC human renal and cardiovascular-associated amino acid sequences of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at: ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 105 AA;

Alignment Scores:
Pred. No.: 6,46-41 Length: 105
Score: 528.00 Matches: 101
Percent Similarity: 95.3% Conservative: 0
Best Local Similarity: 95.3% Mismatches: 4
Query Match: 12.0% Indels: 1
DB: 4 Gaps: 0

US-10-768-158-1 (1-2419) x AAU18674 (1-105)

QY 380 CCTCCCAATGGAGACTCCAGGTCATCTATATGGCTCGCAACCCCAAGGATCTGGTGGT 439
DB 1 ProProGlnTrpArgLeuGlnGlyHisLeuTyrGlySerGlnProGlnGlySerGlyGly 20

QY 440 GTCTTATTATCAGTTCCACCGCTCTCTCGGACCATGAGCTACCGAGGCGACCTTTCAAGA 499
DB 21 ValueLeuSerValProProLeuSerAlaAspHisGluLeuProArg***LeuSerArg 40

QY 500 ATTCTGCGGAGGTTTATGATGATAGCTGGGCTACGGCTCCTCGTTTGGAGCAGTGCA 559
DB 41 IleLeuProGlu-PheMetAsnAspLysLeuGlyTyrGlySerTrpPheGluHisValGI 60

QY 560 GGAGTTCTGGGAGCACCGCATGGACTCGAAGCTGCTTTTCTCAAGTATGAAGACATGCA 619
DB 60 nGluPheTrpGluHisArgMetAspSerAsnValLeuPheLeuLysTyrGluAspMetHi 80

QY 620 TCGGGACCTGGTACGATGGTGGAGCAGCTGGCCAGATTCTCTGGGGGTGTCTCTGTGCAA 679
DB 80 sArgAspLeuValThrMetValGluGlnLeuAlaArgPheLeuGlyValSerCys***I 100

QY 680 GGCCAGCTGGAAGCC 695
DB 100 ePheGlnLeuGluAla 105

RESULT 13
ABU97289
ID ABU97289 standard; protein; 105 AA.
XX AC
XX ABU97289;
XX
DT 30-JUL-2003 (first entry)
DE Human polypeptide #31.
XX
XX Human; kidney disorder; cardiovascular disorder; arrhythmia;
KW glomerulonephritis; urinary tract infection; chronic nephritis; anaemia;
KW carcinoid heart disease; endocarditis; blood disorder; thrombosis;
KW haemoglobin abnormality; electrolyte imbalance; neoplastic disorder;
KW cancer; respiratory disorder; acute rhinitis; sinusitis; pharyngitis;
KW neurological disorder; Alzheimer's disease; Parkinson's disease;
KW Huntington's disease.
XX
OS Homo sapiens.
XX
XX US2003013649-A1.
XX
PD 16-JAN-2003.
XX
XX 21-NOV-2001; 2001US-00989442.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 14-JUL-2000; 2000US-0217496P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.

PR	18-AUG-2000;	2000US-0226279P.	PR	08-NOV-2000;	2000US-0246613P.
PR	22-AUG-2000;	2000US-0226681P.	PR	17-NOV-2000;	2000US-0249207P.
PR	22-AUG-2000;	2000US-0226868P.	PR	17-NOV-2000;	2000US-0249208P.
PR	22-AUG-2000;	2000US-0227182P.	PR	17-NOV-2000;	2000US-0249209P.
PR	23-AUG-2000;	2000US-0227009P.	PR	17-NOV-2000;	2000US-0249210P.
PR	30-AUG-2000;	2000US-0228924P.	PR	17-NOV-2000;	2000US-0249211P.
PR	01-SEP-2000;	2000US-0229287P.	PR	17-NOV-2000;	2000US-0249212P.
PR	01-SEP-2000;	2000US-0229343P.	PR	17-NOV-2000;	2000US-0249213P.
PR	01-SEP-2000;	2000US-0229344P.	PR	17-NOV-2000;	2000US-0249214P.
PR	01-SEP-2000;	2000US-0229345P.	PR	17-NOV-2000;	2000US-0249215P.
PR	05-SEP-2000;	2000US-0229509P.	PR	17-NOV-2000;	2000US-0249216P.
PR	05-SEP-2000;	2000US-0229513P.	PR	17-NOV-2000;	2000US-0249217P.
PR	06-SEP-2000;	2000US-0230437P.	PR	17-NOV-2000;	2000US-0249218P.
PR	06-SEP-2000;	2000US-0230438P.	PR	17-NOV-2000;	2000US-0249244P.
PR	08-SEP-2000;	2000US-0231242P.	PR	17-NOV-2000;	2000US-0249245P.
PR	08-SEP-2000;	2000US-0231243P.	PR	17-NOV-2000;	2000US-0249264P.
PR	08-SEP-2000;	2000US-0231244P.	PR	17-NOV-2000;	2000US-0249265P.
PR	08-SEP-2000;	2000US-0231413P.	PR	17-NOV-2000;	2000US-0249297P.
PR	08-SEP-2000;	2000US-0231414P.	PR	17-NOV-2000;	2000US-0249300P.
PR	08-SEP-2000;	2000US-0232080P.	PR	01-DEC-2000;	2000US-0250160P.
PR	08-SEP-2000;	2000US-0232081P.	PR	01-DEC-2000;	2000US-0250391P.
PR	12-SEP-2000;	2000US-0231968P.	PR	05-DEC-2000;	2000US-0251030P.
PR	14-SEP-2000;	2000US-0232397P.	PR	05-DEC-2000;	2000US-0251988P.
PR	14-SEP-2000;	2000US-0232398P.	PR	05-DEC-2000;	2000US-0256719P.
PR	14-SEP-2000;	2000US-0232399P.	PR	06-DEC-2000;	2000US-0251479P.
PR	14-SEP-2000;	2000US-0232400P.	PR	08-DEC-2000;	2000US-0251856P.
PR	14-SEP-2000;	2000US-0232401P.	PR	08-DEC-2000;	2000US-0251868P.
PR	14-SEP-2000;	2000US-0233063P.	PR	08-DEC-2000;	2000US-0251869P.
PR	14-SEP-2000;	2000US-0233064P.	PR	08-DEC-2000;	2000US-0251989P.
PR	14-SEP-2000;	2000US-0233065P.	PR	08-DEC-2000;	2000US-0251990P.
PR	21-SEP-2000;	2000US-0234223P.	PR	11-DEC-2000;	2000US-0254097P.
PR	21-SEP-2000;	2000US-0234274P.	PR	05-JAN-2001;	2001US-0259678P.
PR	25-SEP-2000;	2000US-0234997P.	PR	17-JAN-2001;	2001US-00764863.
PR	25-SEP-2000;	2000US-0234998P.	XX		
PR	26-SEP-2000;	2000US-0235484P.		(ROSE/) ROSEN C A.	
PR	27-SEP-2000;	2000US-0235834P.	PA	(RUBE/) RUBEN S M.	
PR	27-SEP-2000;	2000US-0235836P.	PA	(BARA/) BARASH S C.	
PR	29-SEP-2000;	2000US-0236327P.	XX		
PR	29-SEP-2000;	2000US-0236367P.	PI	Rosen CA, Ruben SM, Barash SC;	
PR	29-SEP-2000;	2000US-0236368P.	XX	WPI; 2003-416807/39.	
PR	29-SEP-2000;	2000US-0236369P.	DR	N-PSDB; ACD01450.	
PR	29-SEP-2000;	2000US-0236370P.	XX		
PR	02-OCT-2000;	2000US-0236802P.	XX	New nucleic acids and polypeptides, useful for diagnosing, prognosing, preventing and/or treating e.g. kidney, cardiovascular, blood, electrolyte imbalance, neoplastic, respiratory, or neurological diseases or disorders.	
PR	02-OCT-2000;	2000US-0237037P.	PS	Claim 11; Page 275; 363pp; English.	
PR	02-OCT-2000;	2000US-0237038P.	XX		
PR	02-OCT-2000;	2000US-0237039P.	CC	The invention relates to isolated nucleic acids encoding novel polypeptides. The nucleic acids are useful for chromosome mapping, for radiation hybrid mapping, for detection of cancer, in gene therapy, for identifying individuals from minute biological samples, as an alternative to restriction fragment length polymorphism (RFLP) analysis, in forensic biology and as hybridisation probes for differential identification of tissues or cell types present in a biological sample. Compositions comprising the polynucleotides, polypeptides and antibodies specific for the polypeptides may be used in the diagnosis, prognosis, prevention and/or treatment of kidney disorders (e.g. glomerulonephritis, urinary tract infections, chronic nephritis), cardiovascular disorders (e.g. arrhythmias, carinoid heart disease, endocarditis), blood disorders (e.g. thrombosis, anaemia, haemoglobin abnormalities), electrolyte imbalance, neoplastic disorders (e.g. cancers), respiratory disorders (e.g. acute rhinitis, sinusitis, pharyngitis) and neurological disorders (e.g. Alzheimer's disease, Parkinson's disease, Huntington's disease).	
PR	13-OCT-2000;	2000US-0239933P.	CC	Sequences ABU97259-ABU97330 represent human polypeptides of the invention	
PR	20-OCT-2000;	2000US-0240960P.	XX	Sequence 105 AA;	
PR	20-OCT-2000;	2000US-0241221P.	SQ		
PR	20-OCT-2000;	2000US-0241785P.			
PR	20-OCT-2000;	2000US-0241786P.			
PR	20-OCT-2000;	2000US-0241787P.			
PR	20-OCT-2000;	2000US-0241808P.			
PR	20-OCT-2000;	2000US-0241809P.			
PR	20-OCT-2000;	2000US-0241828P.			
PR	01-NOV-2000;	2000US-0244617P.			
PR	08-NOV-2000;	2000US-0246474P.			
PR	08-NOV-2000;	2000US-0246475P.			
PR	08-NOV-2000;	2000US-0246476P.			
PR	08-NOV-2000;	2000US-0246477P.			
PR	08-NOV-2000;	2000US-0246478P.			
PR	08-NOV-2000;	2000US-0246523P.			
PR	08-NOV-2000;	2000US-0246524P.			
PR	08-NOV-2000;	2000US-0246525P.			
PR	08-NOV-2000;	2000US-0246526P.			
PR	08-NOV-2000;	2000US-0246527P.			
PR	08-NOV-2000;	2000US-0246528P.			
PR	08-NOV-2000;	2000US-0246532P.			
PR	08-NOV-2000;	2000US-0246609P.			
PR	08-NOV-2000;	2000US-0246610P.			
PR	08-NOV-2000;	2000US-0246611P.			

PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 06-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231988P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 20-OCT-2000; 2000US-0242221P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.

PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250391P.
PR 01-DEC-2000; 2000US-0251160P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 06-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-541565/60.
DR N-PSDB; ABA12558.
XX
PT Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system cancers
PT and metastases.
XX
XX Claim 11; SEQ ID NO 4889; 1701pp + Sequence Listing; English.
XX
CC The invention relates to novel genes (ABA11004-ABA21534) and proteins
CC (ABA14678-ABB18001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pt_sequences
XX
SQ Sequence 102 AA;

Alignment Scores:
Pred. No.: 4,61e-36 Length: 102
Score: 477.00 Matches: 92
Percent Similarity: 91.1% Conservative: 0
Best Local Similarity: 91.1% Mismatches: 8
Query Match: 10.8% Indels: 1
DB: 4 Gaps: 0

US-10-768-158-1 (1-2419) x ABB16232 (1-102)

Qy 1823 CCGTGGGAGGACCATTTTGGAGGTAGCCCTGTTTCACATCGGATCAGGTGGACGGCC 1882
|||
Db 3 ProLeuGluAspHisPheAlaSerValAlaLeuPheHisSerAspGlnVal-AlaAaGrPr 22
|||
Qy 1883 GCCTGCGTGTCTGTCCACCTCATCCCTCCCTGTATCTGAGGGAGTAAGGTGAGTCTTT 1942
|||

DT	15-JAN-2004 (first entry)	QY	210	GTGCTCTACTTGGTGAGCCAGGGCGCTGACCCCGATGAGATCGGCTTGTGATGAACATCGAC	269
DE	Human disease related protein SeqID111.	Db	57	IIEuAspMetIleTyrGlnGlyGlyAspLeuGlnLysCysAsnArgAlaProIleTyr	76
XX	human; disease state; cytostatic; antiinflammatory; ophthalmological;	QY	270	GAGCAGCTCCCGGTCCTGGAGTACCCACAGCCG-----GGCCTGGACATCATC	317
KW	antiarteriosclerotic; vulnerary; gene therapy;	Db	77	ValArgValProPheLeuGluValAsnAspProGlyGluProSerGlyLeuGluThrLeu	96
KW	hypoxia-regulated condition; tumorigenesis; angiogenesis; apoptosis;	QY	318	AAGAACTGACCTCTCCCGCTCATCAAGAGCCACCTGCCCTACCGCTTCTGCCCTCT	377
KW	inflammation; erythropoiesis; glycolysis; gluconeogenesis;	Db	97	LysAspThrProProArgLeuIleLysSerHisLeuProLeuAlaLeuLeuProGln	116
KW	glucose transportation; catecholamine synthesis; iron transport;	QY	378	GACCTCCACAATGGAGACTCCCAAGTCTATATGCTCGCAACCCCAAGATCTGGTG	437
KW	nitric oxide synthesis; cancer; ischaemic condition; reperfusion injury;	Db	117	ThrLeuLeuAspGlnLysValLysValValAlaArgAsnProLysAspValAla	136
KW	retinopathy; neonatal stress; pre-eclampsia; atherosclerosis;	QY	438	GTGTCTTATTATCAGTTCCACCGCTCTCTGGGACCATGAGCTACCGAGGCACCTTTCAA	497
XX	inflammatory condition; wound healing.	Db	137	ValSerTyrTyrHisPheHisArgMetGluLysAlaHisProGluProGlyThrTrpAsp	156
OS	Homo sapiens.	QY	498	GAATTCTCGCGAGGTTTATGATGATGAATGATGAAGCTGGGCTACGGCTCTGTTTGACACGTG	557
XX	WO2003018621-A2.	Db	157	SerPheLeuGluLysPheMetAlaGlyGluValSerTyrGlySerTyrGlnHisVal	176
XX	06-MAR-2003.	QY	558	CAGGAGTTCTGGGAGCACCGCATGGACTCGAAACGTGCTTTTCTCAAGTATGAAGACATG	617
XX	23-AUG-2002; 2002WO-GB003892.	Db	177	GlnGluTrpGluLeuSerArgThrHisProValLeuTyrLeuPheTyrGluAspMet	196
XX	23-AUG-2001; 2001GB-00020558.	QY	618	CATCGGACCTGGTGACGATGGTGAGCAGCTGGCCAGATTCTCTGGGGTGCTCTGTGAC	677
PR	05-OCT-2001; 2001GB-00024037.	Db	197	LysGluAsnProLysArgGluIleGlnLysIleLeuGluPheValGlyArgSerLeuPro	216
XX	(OXFO-) OXFORD BIOMEDICA UK LTD.	QY	678	AAGCCACGATGGAGCCCTGACGGACACATGC-----	710
XX	Kingsman SM, White J, Ward NR, Harris RA, Naylor S, Mundy CR;	Db	217	GluGluThrMetAspPheMetValGlnHisThrSerPheLysGluMetLysLysAsnPro	236
DR	WPI; 2003-290046/28.	QY	711	-----CACAGCTGGTGACCAAGCTGCTGCAACGCTGAGGCC	746
DR	N-PSDB; ADD18681.	Db	237	MetThrAsnTyrThrValProGlnGluLeuMetAspHisSerIleSer-----	253
XX	New substantially purified polypeptide, useful for diagnosing or treating	QY	747	CTGCCC---GTGGCCCGGGAAGAGTTGGGCTGTGGAAGGACATCTTTCACCGTCTCCATG	803
PT	a hypoxia-regulated condition, such as cancer, ischemia, reperfusion	Db	254	---ProPheMetArgLysGlyMetAlaGlyAspTrpLysThrThrPheThrValAlaGln	272
PT	injury, retinopathy, pre-eclampsia, atherosclerosis, inflammation, or	QY	804	AATCAGAGTTGACTTGGTGATATAACACAGAGATGGAAAGTGTGACCTCACGTTT	860
XX	wound healing.	Db	273	AsnGluArgPheAspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe	291
XX	Claim 25; SEQ ID NO 111; 424pp; English.	RESULT 19			
XX	This invention relates to novel human genes and gene product which are	ADP24302			
CC	implicated in certain disease states. Compounds which modulate the	ID	ADP24302	standard; protein; 295 AA.	
CC	proteins of the invention may have cytostatic, antiinflammatory,	XX	AC	ADP24302;	
CC	ophthalmological, antiarteriosclerotic or vulnerary activities. The	XX	DT	18-NOV-2004 (first entry)	
CC	sequences of the invention may be useful for gene therapy. The invention	XX	DE	PRO polypeptide SEQ ID NO:1480.	
CC	may be useful for diagnosing or treating a hypoxia-regulated condition,	XX	KW	PRO; antiinflammatory; antiarthritic; antirheumatic; immunosuppressive;	
CC	such as tumorigenesis, angiogenesis, apoptosis, inflammation,	XX	KW	osteopathic; antidiabetic; dermatological; antipsoriatic; antiallergic;	
CC	erythropoiesis, or the biological response to hypoxia conditions	XX	KW	antiasthmatic; hepatotropic; respiratory; gene therapy; immune system.	
CC	including processes such as glycolysis, gluconeogenesis, glucose	XX	XX	Unidentified.	
CC	transportation, catecholamine synthesis, iron transport or nitric oxide	XX	XX	WO2004041170-A2.	
CC	synthesis. The disease includes cancer, ischaemic conditions, reperfusion	XX	XX	21-MAY-2004.	
CC	injury, retinopathy, neonatal stress, pre-eclampsia, atherosclerosis,	XX	XX	30-OCT-2003; 2003WO-US034312.	
CC	inflammatory conditions or wound healing. The present sequence is that of	XX	XX	01-NOV-2002; 2002US-0423394P.	
XX	a disease related protein of the invention.	XX	PA	(GETH) GENENTECH INC.	
SQ	Sequence 295 AA;				

Alignment Scores:
 Pred. No.: 2,43e-35 Length: 295
 Score: 472.00 Matches: 100
 Percent Similarity: 54.5% Conservative: 52
 Best Local Similarity: 35.8% Mismatches: 101
 Query Match: 10.7% Indels: 26
 DB: 7 Gaps: 4

US-10-768-158-1 (1-2419) x ADD18680 (1-295)

QY	90	GGCGTGGCTGCCCGCTTCTGCGCGGGAAGATGGAGAGATCGCAACTTCCCGGTG	149
Db	17	GlyValProLeuIleLysTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla	36
QY	150	CGGCCACGACGCTGGATGCTACCTACCCCAAGTCCGGACACGATCTGCTGAGGAG	209
Db	37	ArgProAspAspLeuLeuIleAsnThrTyrProLysSerGlyThrThrTrpValSerGln	56

XX Genetically typing animals to determine those with desired boar taint
 PT characteristics, comprises obtaining a sample of genetic material from
 PT the animal, and assaying for the presence of a sulfotransferase allele.
 XX
 PS Example; SEQ ID NO 4; 24pp; English.
 XX
 CC The present invention relates to a method of genetically typing animals
 CC to determine those with desired boar taint characteristics. The method
 CC involves obtaining a sample of genetic material from the animal and
 CC assaying for the presence of a sulfotransferase allele. The present
 CC sequence is the human phenol sulfotransferase (PST) protein, SULT1A3. The
 CC SULT1A3 protein encoding gene is located on chromosome 16p12.1.
 XX
 SQ Sequence 295 AA;

Alignment Scores:
 Pred. No.: 2,43e-35 Length: 295
 Score: 472.00 Matches: 100
 Percent Similarity: 54.5% Conservative: 52
 Best Local Similarity: 35.8% Mismatches: 101
 Query Match: 10.7% Indels: 26
 DB: 9 Gaps: 4

US-10-768-158-1 (1-2419) x ADM71803 (1-295)

QY 90 GCGTGGCGTGGCGCCCTTCTGCGGGGGAAGATGGAGAGATCGGCAACTTCCCGGTG 149
 DB 17 GlyValProLeuIleLysTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla 36
 QY 150 CGGCCACGACGTGTGATCGTACCTACCCAGTCCGCGCAGCTTCTGCTGCGAGG 209
 DB 37 ArgProAspLeuLeuIleAsnThrTyrProLysSerGlyThrTrpTrpValSerGln 56
 QY 210 GTGTCTACTTGGTGGCGGCGCTGACCCCGATGAGATCGGCTTGATGAACATCGAC 269
 DB 57 IleLeuAspMetIleTyrGlnGlyAspLeuGluLysCysAsnArgAlaProIleTyr 76
 QY 270 GAGCAGTCCCGTCTCGGATACCCACGCGG-----GGCTGGACATCATC 317
 DB 77 ValArgValProPheLeuGluValAsnAspProGlyGluProSerGlyLeuGluThrLeu 96
 QY 318 AAGGAACGTACTCTCCCGCTCATCAAGAGCCACCTGCCCTACCGCTTTCGCCCTCT 377
 DB 97 LysAspThrProProArgGluLeuIleLysSerHisLeuProLeuAlaLeuProGln 116
 QY 378 GACTCCCAATGGAGACTCAAGGTCTATATGGCTCGCAACCCCAAGGATCTGGTG 437
 DB 117 ThrLeuLeuAspGlnLysValLysValValTyrValAlaArgAsnProLysAspValAla 136
 QY 438 GTGCTTATTATCAGTTCCACGGCTCTGCGGACCATTGAGTACCGAGGACCTTTCAA 497
 DB 137 ValSerTyrTyrHisPheHisArgMetGluLysAlaHisProGluProGlyThrTrpAsp 156
 QY 498 GAATTCCTCCGAGGTTTGAATGATGAAGCTGGCTACGGCTCCTGTTGGACACGTG 557
 DB 157 SerPheLeuGluLysPheMeAlaGlyGluValSerTyrGlySerTyrGlnHisVal 176
 QY 558 CAGGAGTCTCGGAGCACCGCATGACTCGAACGTGCTTTTCTCAAGTATCAAGACATG 617
 DB 177 GlnGluTrpTrpGluLeuSerArgThrHisProValLeuTyrLeuPheTyrGluAspMet 196
 QY 618 CATCGGACCTGGTGACGATGGTGAGCAGTGGCGCAGATTCCTGGGGGTCTCCTGTGAC 677
 DB 197 LysGluAsnProLysArgGluIleGlnLysIleLeuGluPheValGlyArgSerLeuPro 216
 QY 678 AAGGCCAGCTGGAAGCCCTGACGAGCAGCTGC----- 710
 DB 217 GluThrMetAspPheMetValGlnHisThrSerPheLysGluMetLysLysAsnPro 236
 QY 711 -----CACCAGCTGGTGACCACTGCTGCAACGCTGAGGCC 746
 DB 237 MetThrAsnTyrThrValProGlnGluLeuMetAspHisSerIleSer----- 253

QY 747 CTGCC---GTGGCGCGGGAAGAGTTGGCTGTGAAGGACATCTTTCACGCTCTCCATG 803
 DB 254 ---ProPheMetArgLysGlyMetAlaGlyAspTrpLysThrThrPheThrValAlaGln 272
 QY 804 AATCAGAAGTTTACCTTGGTGTTATAACACAGAAGATGGAAAAGTGTGTGACCTCAGCTTT 860
 DB 273 AsnGluArgPheAspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe 291
 RESULT 21
 ADI17130
 ID ADI17130 standard; protein; 296 AA.
 XX
 AC ADI17130;
 XX
 DT 15-APR-2004 (first entry)
 XX
 DE Rat NOVX protein homologue SeqID 666.
 XX
 KW rat; NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes;
 KW inflammation; autoimmune disorder; allergy; blood disorder;
 KW acquired immunodeficiency syndrome; AIDS; obesity; asthma;
 KW immunoglobulin (Ig)A nephropathy; cirrhosis; arthritis;
 KW Alzheimer's disease; infection; str.
 XX
 OS Rattus norvegicus.
 XX
 FN WO200268649-A2.
 XX
 PD 06-SEP-2002.
 XX
 PF 31-JAN-2002; 2002WO-US002785.
 XX
 PR 31-JAN-2001; 2001US-0265395P.
 PR 31-JAN-2001; 2001US-0265412P.
 PR 31-JAN-2001; 2001US-0265514P.
 PR 31-JAN-2001; 2001US-0265517P.
 PR 02-FEB-2001; 2001US-0266406P.
 PR 05-FEB-2001; 2001US-0266787P.
 PR 07-FEB-2001; 2001US-0266975P.
 PR 08-FEB-2001; 2001US-0267057P.
 PR 09-FEB-2001; 2001US-0267459P.
 PR 15-FEB-2001; 2001US-0268974P.
 PR 26-FEB-2001; 2001US-0271664P.
 PR 27-FEB-2001; 2001US-0271839P.
 PR 27-FEB-2001; 2001US-0271855P.
 PR 02-MAR-2001; 2001US-0273046P.
 PR 14-MAR-2001; 2001US-0275925P.
 PR 14-MAR-2001; 2001US-0275947P.
 PR 14-MAR-2001; 2001US-0275950P.
 PR 15-MAR-2001; 2001US-0276448P.
 PR 16-MAR-2001; 2001US-0276450P.
 PR 16-MAR-2001; 2001US-0276397P.
 PR 20-MAR-2001; 2001US-0276768P.
 PR 26-MAR-2001; 2001US-0278652P.
 PR 26-MAR-2001; 2001US-0278775P.
 PR 29-MAR-2001; 2001US-0278882P.
 PR 29-MAR-2001; 2001US-0279884P.
 PR 30-MAR-2001; 2001US-0280147P.
 PR 11-APR-2001; 2001US-0282992P.
 PR 11-APR-2001; 2001US-0283083P.
 PR 20-APR-2001; 2001US-0285133P.
 PR 23-APR-2001; 2001US-0285749P.
 PR 03-MAY-2001; 2001US-0288327P.
 PR 03-MAY-2001; 2001US-0288504P.
 PR 29-MAY-2001; 2001US-0294047P.
 PR 30-MAY-2001; 2001US-0294473P.
 PR 08-JUN-2001; 2001US-0296964P.
 PR 18-JUN-2001; 2001US-0298959P.

PR 19-JUN-2001; 2001US-0299324P.
 PR 13-AUG-2001; 2001US-0312020P.
 PR 16-AUG-2001; 2001US-0312889P.
 PR 16-AUG-2001; 2001US-0312908P.
 PR 21-AUG-2001; 2001US-0313390P.
 PR 28-AUG-2001; 2001US-0315470P.
 PR 31-AUG-2001; 2001US-0316447P.
 PR 07-SEP-2001; 2001US-0318115P.
 PR 12-SEP-2001; 2001US-0318118P.
 PR 19-SEP-2001; 2001US-0323379P.
 PR 18-OCT-2001; 2001US-0330245P.
 PR 18-OCT-2001; 2001US-0330308P.
 PR 14-NOV-2001; 2001US-0332701P.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Tchernev VT, Spyttek KA, Zerhusen BD, Patturajan M, Shinkets RA;
 PI Li L, Gangolli EA, Padigaru M, Anderson DW, Rastelli L, Miller CE;
 PI Gerlach VL, Taupier RU, Gusev VY, Colman SD, Wolenc AR, Pena CE;
 PI Furtak K, Grosse WM, Alabrook JP, Lepley DM, Rieger DK, Burgess CE;
 XX
 DX WPI; 2002-706998/76.
 XX
 DX New NOVX polypeptides and nucleic acids, useful for preventing or
 PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
 PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
 PT pharmacogenomics.
 XX
 PS Disclosure; SEQ ID NO 666; 1498pp; English.
 XX
 CC This invention relates to a novel nucleic acids, and encoded polypeptides
 CC thereof, which have properties related to the stimulation of biochemical
 CC or physiological responses in a cell, tissue, organ or organism.
 CC Specifically, it refers to the use of biologically active fragments for
 CC diagnostic and prognostic assays and furthermore in the treatment of
 CC diverse pathological conditions. The present invention describes novel
 CC human and murine NOVX proteins, as well as methods to modulate their
 CC expression using antisense oligos, ribozymes and peptide nucleic acids.
 CC The NOVX polypeptides, polynucleotides and antibodies are useful in
 CC treating or preventing NOVX-associated disorders, e.g. cardiomyopathy,
 CC atherosclerosis, cancer and diabetes. Furthermore, they may be used in
 CC treating or preventing diseases such as inflammation, autoimmune
 CC disorders, allergies, blood disorders, acquired immunodeficiency syndrome
 CC (AIDS), obesity, asthma, immunoglobulin (IgA) nephropathy, cirrhosis,
 CC arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy
 CC and epilepsy. Accordingly, these molecules have many activities including
 CC cyostatic, cardiac, antiinflammatory, immunosuppressive, antiallergic,
 CC haemostatic, anti-HIV, antidiabetic, antiarteriosclerotic, anorectic,
 CC antiasthmatic, nephrotropic, antiarthritic, virucide, antiparasitic,
 CC neuroprotective, nootropic, antibacterial, virucide, antiparasitic,
 CC relaxant and anticonvulsant. In addition, they are useful in screening
 CC assays to identify small molecules that modulate or inhibit, for example,
 CC neurogenesis, wound healing and angiogenesis. The nucleic acids are also
 CC used as in chromosome mapping, tissue typing, preventive medicine and
 CC pharmacogenomics. This polypeptide is a homologue of a human NOVX protein
 CC of the invention.
 XX
 SQ Sequence 296 AA;
 Alignment Scores:
 Pred. No.: 7.29e-35 Length: 296
 Score: 467.00 Matches: 100
 Percent Similarity: 55.4% Conservative: 58
 Best Local Similarity: 35.1% Mismatches: 105
 Query Match: 10.6% Indels: 22
 DB: 5 Gaps: 5
 US-10-768-158-1 (1-2419) x ADI17130 (1-296)
 QY 66 GAGAGCAAGTACTTCGAGTTCCTGCGGTGCGGCTGCGGCTTCTCTGCGGGAGATG 125
 : : : : :
 Db 10 GlnThrLysLeuLysGluValAlaGlyLeuProLeuGlnAlaProThrValAspAsnTrp 29

QY 126 GAGAGATCCCAACTTCCCGTGCAGCCAGCAGCTGTGGATCGTCACCTACCCCAAG 185
 : : : : :
 Db 30 SerGlnIleGlnThrPheLysAlaLysProAspAspLeuLeuIleCysThrTrpLys 49
 : : : : :
 QY 186 TCGGACACAGCTGTCTGAGGAGGTCTTACTTGTGTGAGCCAGGCGCTGACCCGAT 245
 : : : : :
 Db 50 SerGlyThrThrTrpIleGlnGluLeuValAspMetIleGluGlnAsnGlyAspValGlu 69
 : : : : :
 QY 246 GAGATCGGCTTGATGAACATCGACGAGCAGCTCCCGCTCTGGAGTAC-----CCA 296
 : : : : :
 Db 70 LysCysGlnArgThrIleIleGlnHisArgHisProPheIleGluTrpAlaArgProPro 89
 : : : : :
 QY 297 CAGCG---GGCTGGACATCATCAAGAACTGACCTCTCCCGCTCATCAAGAGCAC 353
 : : : : :
 Db 90 GlnProSerGlyValAspLysAlaAsnAlaMetProAlaProArgIleLeuArgThrHis 109
 : : : : :
 QY 354 CTGCCCTACCGCTTCTGCGCTCTGACCTCCACATCGAGACTCCAGGTCATCATATG 413
 : : : : :
 Db 110 LeuProThrGlnLeuLeuProProSerPheTrpThrAsnAsnCysLysPheLeuTrpVal 129
 : : : : :
 QY 414 GCTCGCAACCCCAAGGATCTGTGTCTTATTATCAGTTTCCACCGCTCTCTCGGACC 473
 : : : : :
 Db 130 AlaArgAsnAlaLysAspCysMetValSerTrpHisPheTrpArgMetSerGlnVal 149
 : : : : :
 QY 474 ATGAGCTACCGAGCACCTTCAAGAACTCTCGCGGAGTTTATGAATGATTAAGCTGGC 533
 : : : : :
 Db 150 LeuProAspProGlyThrTrpAsnGluTrpPheGluThrPheIleAsnGlyLysValSer 169
 : : : : :
 QY 534 TAGGCTCTCGTGTGAGCAGCTGCGAGGCTTCTGGGAGCACCGCATGGAGCTCGAACTG 593
 : : : : :
 Db 170 TrpGlySerTrpPheAspHisValLysGlyTrpTrpGluLeuArgAspArgTyrGlnIle 189
 : : : : :
 QY 594 CTTTTTCTCAAGTATGAAGACATCGATCGGAGCTGTGTGACGATGTGGAGCAGCTGGCC 653
 : : : : :
 Db 190 LeuPheLeuPheTyrGluAspValLysArgAspProLysArgGluIleGlnLysValMet 209
 : : : : :
 QY 654 AGATTCTCGGGGTGTC-----TGTGACAAAGGCCAGCTGGAAGCCCTG 698
 : : : : :
 Db 210 GlnPheMetGlyLysAsnLeuAspGluGluValValAspLysIleValLeuGluThrSer 229
 : : : : :
 QY 699 ACGAGCAGCTGCCACCGCTGTGTGGACGCTGTGCAACGCTGAGGCGCTGCC----- 752
 : : : : :
 Db 230 PheGluLysMetLysGlu-----AsnProMetThrAsnArgSerThrValProLysSer 247
 : : : : :
 QY 753 -----GTGGCGGGGAGAGTTGGGCTGTGGAAGGAC 785
 : : : : :
 Db 248 ValLeuAspGlnSerIleSerProPheMetArgLysGlyThrValGlyAspTrpLysAsn 267
 : : : : :
 QY 786 ATCTTCACCTCTCCATGAATGAGAGTTTGACTTGTATTAACAGAGATCGGAAG 845
 : : : : :
 Db 268 HisPheThrValAlaGlnAsnAspArgPheAspGluIleTyrLysGlnLysMetGlyGly 287
 : : : : :
 QY 846 TGTGACCTCAGTTT 860
 : : : : :
 Db 288 ThrSerLeuAsnPhe 292
 : : : : :
 RESULT 22
 ADY80433
 ID ADY80433 standard; protein; 285 AA.
 XX
 AC ADY80433;
 XX
 DT 19-MAY-2005 (first entry)
 XX
 DE Rat aryl sulfotransferase, SEQ ID 1.
 XX
 KW diagnosis; diagnostic; dioxin exposure; aryl sulfotransferase.
 XX
 OS Rattus norvegicus.
 XX
 PN KR2004062834-A.
 XX

PD 09-JUL-2004.
XX 03-JAN-2003; 2003KR-00000367.
XX 03-JAN-2003; 2003KR-00000367.
XX (MITE-) MI TECH CO LTD.
XX
XX Kim CW, Lee DY, Son WG;
XX WPI; 2004-763526/75.
XX Biomarker proteins for diagnosing the exposure to dioxin.
XX Example 3; SEQ ID NO 1; lipp; Korean.
XX
XX The present invention relates to biomarker proteins for diagnosing
CC exposure to dioxin. The biomarker proteins are also useful in development
CC of protein chips detecting hard-decomposed pollutants. The biomarker
CC proteins of the invention comprise aryl sulfoxtransferase (ADY80433) which
CC is present only in the blood of an individual exposed to dioxin and has
CC a molecular weight of 26 kDa and pI 5.7, or a protein which is expressed at
CC a higher level in the blood of an individual exposed to dioxin than in
CC that of the normal individual, and is selected from Cytochrome 8
CC polypeptide having molecular weight of 52 kDa and pI 5.5, Calcium-
CC dependent protease having molecular weight of 28 kDa and pI 5.3 and
CC Proteasome subunit beta type 4 precursor having molecular weight of 29
CC kDa and pI 6.4.
XX
XX Sequence 285 AA;
SQ
Alignment Scores:
Pred. No.: 5.15e-34 Length: 285
Score: 458.00 Matches: 96
Percent Similarity: 55.6% Conservative: 57
Best Local Similarity: 34.9% Mismatches: 104
Query Match: 10.4% Indels: 18
DB: 8 Gaps: 4
US-10-768-158-1 (1-2419) x ADY80433 (1-285)
QY 90 GGCGTGGCGTCCGCGCTTCTGCGCGGGAAGATGGAGATCGGCACTTCCCGGTG 149
DB 7 GlyLeuProLeuIleuLysTyrPheAlaGluThrIleGlyProLeuGlnAnPheThrAla 26
QY 150 CGGCCAGCGACGTGTGGATGTCACCTACCCCAAGTCCGCGACACGCTTGCTGCAGAG 209
DB 27 TrpProAspAepLeuLeuIleSerThrTyrProLysSerGlyThrThrTrpMetSerGlu 46
QY 210 GTGCTCTACTTGGTGACCGCGGCTGACCCCGATGAGATCGGCTTGATGAACATCGAC 269
DB 47 IleLeuAspMetIleTyrGlnGlyGlyLeuGlnLysCysGlyArgAlaProIleTyr 66
QY 270 GAGCAGCTCCCGTCTCGGAGTACCCACAGCCG-----GGCCTGGACATCATC 317
DB 67 AlaArgValProPheLeuGluPheLysCysProGlyValProSerGlyLeuGluThrLeu 86
QY 318 AAGAACTGACCTCTCCCGGCTCATCAAGAGCCACCTGCGCTTACCCTTTTGGCCCTCT 377
DB 87 GluGluThrProAlaProArgLeuLeuLysThrHisLeuProLeuSerLeuLeuProGln 106
QY 378 GACCTCCACATGAGACTCCAGTCTATATGCTCGACCTCGACCCCAAGGATCTGGTG 437
DB 107 SerLeuLeuAspGlnLysValLysValIleTyrIleAlaArgAsnAlaLysAspValVal 126
QY 438 GTGTCTTATTATCATGTTCCACCGCTCTCTGCGGACCATGAGCTACCGAGGACCTTTCAA 497
DB 127 ValSerTyrTyrAsnPheTyrAsnMetAlaLysLeuHisProAspProGlyThrTrpAsp 146
QY 498 GAATTCGCGGAGGTTTATGAATGATGAAGCTGGGCTACGGCTCTCTGGTTGACGAGTGC 557
DB 147 SerPheLeuGluAanPheMetAspGlyGluValSerTyrGlySerTyrTrpTrpGlnHisVal 166

QY 558 CAGGAGTTCGGGAGCACCAGCATGGACTCGAACGTCGCTTTTCTCAAGTATGAAGCATG 617
DB 167 LysGluTrpTrpGlnLeuArgHisThrHisProValLeuTyrLeuPheTyrGluAspIle 186
QY 618 CATCGGACCTGCTGACGATGGTGGAGCAGCTGGCCAGATTCCTGGGGGTGTCTGTGAC 677
DB 187 LysGluAsnProLysArgGluIleLysLysIleLeuGluPheLeuGlyArgSerLeuPro 206
QY 678 AAGGCCAGCTGGAAGCCCTGACGGAGCACTGC-----CACCAGCTGGTGGACCATGTC 731
DB 207 GluGluThrValAspSerIleValHisHisThrSerPheLysLysMetLysGluAsnCys 226
QY 732 ---TGCAACGCTGAGGCCCTGCC-----GTG 755
DB 227 MetThrAsnTyrThrThrIleProThrGluIleMetAspHisAsnValSerProPheMet 246
QY 756 GGCGCGGGAAGATGGGCTGTGGAAGGACATCTTACCCGCTCCCATGATGAGAATTT 815
DB 247 ArgLysGlyThrThrGlyAspTrpLysAsnThrPheThrValAlaGlnAsnGluArgPhe 266
QY 816 GACTTGGTGTATAAACAGAAAGATGGGAAAGTGTGACCTCAGCTT 860
DB 267 AspAlaHisTyrAlaLysThrMetThrAspCysAspPheLysPhe 281
RESULT 23
ADD45448
ID ADD45448 standard; protein; 291 AA.
XX
XX ADD45448;
XX
XX 29-JAN-2004 (first entry)
XX
XX Rat Protein P17988, SEQ ID NO 10881.
XX
XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
XX chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
XX Rattus norvegicus.
XX
XX WO2003016475-A2.
XX
XX 27-FEB-2003.
XX
XX 14-AUG-2002; 2002WO-US025765.
XX
XX 14-AUG-2001; 2001US-0312147P.
XX
XX 01-NOV-2001; 2001US-0346382P.
XX
XX 26-NOV-2001; 2001US-0333347P.
XX
XX (GEHO) GEN HOSPITAL CORP.
XX
XX (FARB) BAYER AG.
XX
XX Woolf C, D'urso D, Befort K, Costigan M;
XX
XX WPI; 2003-268312/26.
XX
XX GENBANK; P17988.
XX
XX New composition comprising two or more isolated polypeptides, useful for
XX preparing a medicament for treating pain in an animal.
XX
XX Claim 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
XX or human polynucleotides or a polynucleotide which represents a fragment,
XX derivative or allelic variation of the nucleic acid sequence. Also
XX claimed are a vector comprising the novel polynucleotide, a host cell
XX comprising the vector, a method for identifying a nucleotide sequence
XX which is differentially regulated in an animal subjected to pain and a
XX kit to perform the method, an array, a method for identifying an agent
XX that increases or decreases the expression of the polynucleotide sequence
XX that is differentially expressed in neuronal tissue of a first animal
XX subjected to pain, a method for identifying a compound which regulates
XX the expression of a polynucleotide sequence which is differentially

CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 291 AA;

Alignment Scores:
Pred. No.: 5.21e-34 Length: 291
Score: 458.00 Matches: 96
Percent Similarity: 55.6% Conservative: 57
Best Local Similarity: 34.9% Mismatches: 104
Query Match: 10.4% Indels: 18
DB: 7 Gaps: 4

US-10-768-158-1 (1-2419) x ADD45448 (1-291)

Qy 90 GGCGTGGCGTCCGCGCTTCTGCGCGGGAAGATGGAGGATCGCCCACTTCCCGTG 149
Db 13 GlylleProleuileLysTyPheAlaGluThrIleGlyProLeuGlnAsnPheThAla 32
Qy 150 CGGCCCGCAGCGTGTGGATCGTCACCTACCCCAAGTCCGGCAGCTTCTCGCAGAG 209
Db 33 TrpProaspheLeuLeuSerThrTyProlysserGlyThrTrpMetSerGlu 52
Qy 210 GTGTCTACTTGGTGGCAGGCGCTGACCCCGATGATGATCGGCTGTGATGAACATGC 269
Db 53 ileLeuaspMetIleTyGlnGlyGlyLeuGluLysCysGlyArgAlaProIleTy 72
Qy 270 GAGCAGCTCCCGTCTGGAGTACCCACAGCCG-----GGCCTGGACATCATC 317
Db 73 AlaargvalProPheLeuGluPheLysCysProGlyValProSerGlyLeuGluThrLeu 92
Qy 318 AAGAACTGACCTCTCCCGCTCATCAAGAGCCACCTGCCTACCGCTTCTGCCCTCT 377
Db 93 GluGluThrProAlaProArgLeuLeuLysThrHisLeuProLeuSerLeuLeuProGln 112
Qy 378 GACCTCCACATGAGACTCCAAAGTCAATCATATATGCTCGCAACCCCAAGATCTGGT 437
Db 113 SerLeuLeuaspGlnLysValLysValIleTyIleAlaArgAsnAlaLysAspValVal 132
Qy 438 GTGTCTTATTATCATGTTCCACCGCTCTCTCGGACCATGAGTACCGAGCCACCTTCAA 497
Db 133 ValserTyTyAsnPheTyAsnMetAlaLysLeuHisProAspProGlyThrTrpAsp 152
Qy 498 GAATTCCTCCCGAGGTTTATGAATGATAAGCTGGCTACGGCTCTGTGTTGACACGTG 557
Db 153 SerPheLeuGluAsnPheMetAspGlyValSerTyGlySerTyGlyThrGlnHisVal 172
Qy 558 CAGGAGTTCTGGGAGCACCAGCATGAGTCAACCTGTTTCTCAAGTATGAAGACATG 617
Db 173 LysGluThrPrpGluLeuArgHisThrHisProValLeuTyLeuPheTyGluAspIle 192
Qy 618 CATCGGACCTGTGACATGTTGGAGCAGCTGCCAGATTCCTGGGGGTGCTCTGTGAC 677
Db 193 LysGluAsnProLysArgGluIleLysLeuGluLeuGluPheLeuGlyArgSerLeuPro 212
Qy 678 AAGGCCCGAGCTGGAGCCCTGACGAGCACTGC-----CACCAAGCTGGTGGACCACTGC 731
Db 213 GluGluThrValaspSerIleValHisThrSerPheLysLysMetLysGluAsnCys 232

Qy 732 ---TGCACCGCTGAGGCCCTGCCCC-----GTG 755
Db 233 MetThrAsnTyThrThrIleProThrGluIleMetAspHisAsnValSerProPheMet 252
Qy 756 GGCCGGGGAAGAGTTGGCTGTGGAAGGACATCTTACCGTCTCATGATGAAGAAGTTT 815
Db 253 ArgLysGlyThrThrGlyAspTrpLysAsnThrPheThrValAlaGlnAsnGluArgPhe 272
Qy 816 GACTTGGTGTATAACAGAGATCGGAAAGTGTGACCTCAGCTT 860
Db 273 AspAlaHisTyAlaLysThrMetThrAspCysAspPheLysPhe 287
RESULT 24
ADE57139
ID ADE57139 standard; protein; 291 AA.
XX ADE57139;
AC ADE57139;
DT 29-JAN-2004 (first entry)
XX Rat Protein P17988, SEQ ID NO 2999.
DE Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
XX chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
OS Rattus norvegicus.
XX WO2003016475-A2.
XX 27-FEB-2003.
XX 14-AUG-2002; 2002WO-US025765.
XX 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX (GEHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX Woolf C, D'urso D, Befort K, Costigan M;
XX WPI; 2003-268312/26.
DR GENBANK; P17988.
XX New composition comprising two or more isolated polypeptides, useful for
XX preparing a medicament for treating pain in an animal.
PS Claim 1; Page; 1017pp; English.
XX The invention discloses a composition comprising two or more isolated rat
XX or human polynucleotides or a polynucleotide which represents a fragment,
XX derivative or allelic variation of the nucleic acid sequence. Also
XX claimed are a vector comprising the novel polynucleotide, a host cell
XX comprising the vector, a method for identifying a nucleotide sequence
XX which is differentially regulated in an animal subjected to pain and a
XX kit to perform the method, an array, a method for identifying an agent
XX that increases or decreases the expression of the polynucleotide sequence
XX that is differentially expressed in neuronal tissue of a first animal
XX subjected to pain, a method for identifying a compound which regulates
XX the expression of a polynucleotide sequence which is differentially
XX expressed in an animal subjected to pain, a method for identifying a
XX compound that regulates the activity of one or more of the
XX polynucleotides, a method for producing a pharmaceutical composition, a
XX method for identifying a compound or small molecule that regulates the
XX activity in an animal of one or more of the polypeptides given in the
XX specification, a method for identifying a compound useful in treating
XX pain and a pharmaceutical composition comprising the one or more
XX polypeptides or their antibodies. The polynucleotide or the compound that
XX modulates its activity is useful for preparing a medicament for treating
XX pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
XX injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
XX therapy). The sequence presented is a rat protein (shown in Table 2 of

CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 291 AA;

Alignment Scores:
Pred. No.: 5.21e-34 Length: 291
Score: 458.00 Matches: 96
Percent Similarity: 55.6% Conservative: 57
Best Local Similarity: 34.9% Mismatches: 104
Query Match: 10.4% Indels: 18
DB: 7 Gaps: 4

US-10-768-158-1 (1-2419) x ADE57139 (1-291)

```
QY 90 GCGCTGCGGCTGCGCGCTCTCGCGGGGAAGATGGAGGATCGCCAACTTCCCGTG 149
D 13 GlylleProLeuIleLysTyrPheAlaGluThrIleGlyProLeuGlnAsnPheThrAla 32
QY 150 CGGCCCGACGCGTGTGGATCGTCACTACCCCAAGTCCGCCACGAGCTTGTCCGAGG 209
D 33 TrpProAspAspLeuLeuIleSerThrTyrProLysSerGlyThrTrpMetSerGlu 52
QY 210 GTGTCTACTTGGTGACCGGCGCTGACCCCGATGAGATCGGCTTGATGAACATCGAC 269
D 53 IleLeuAspMetIleTyrGlnGlyLysLeuGluLysCysGlyArgAlaProIleTyr 72
QY 270 GAGCAGCTCCCGGCTCGGATGACCCACAGCGG-----GGCTGGACATCATC 317
D 73 AlaArgValProPheLeuGluPheLysCysProGlyValProSerGlyLeuGluThrLeu 92
QY 318 AAGAACTGACCTTCCCGCTCATCAAGACCCACTGCGCTTCTGCGCTCT 377
D 93 GluGluThrProAlaProArgLeuLeuLysThrHisLeuProLeuSerLeuProGln 112
QY 378 GACCTCCCAATGAGAGTCCAGGTCTATATGGCTCGCAACCCCAAGATCGTG 437
D 113 SerLeuLeuAspGlnLysValIleTyrIleAlaArgAsnAlaLysAspValVal 132
QY 438 GTGTCTTATTATGTTCCACCGCTCTCTCGGACCATGAGTACCGAGGACCTTCAA 497
D 133 ValSerTyrTyrAsnPheTyrAsnMetAlaLysLeuHisProAspProGlyThrTrpAsp 152
QY 498 GAATTTCTCGCGAGGTTTATGAATGATTAAGCTGGCTACGGCTCTGTTGAGCAGCTG 557
D 153 SerPheLeuGluAsnPheMetAspGlyGluValSerTyrGlySerTyrGlnHisVal 172
QY 558 CAGGAGTCTGGAGCACCAGTACGACTCGAACGCTTTTCTCAAGTATGAGACATG 617
D 173 LysGluThrTrpGluLeuArgHisThrHisProValLeuTyrLeuPheTyrGluAspIle 192
QY 618 CATCGGACCTGGTGACCATGTGGAGCAGCTGGCCAGATTCCTGGGGGTCTCTGTGAC 677
D 193 LysGluAsnProLysArgGluIleLysLysIleLeuGluPheLeuGlyArgSerLeuPro 212
QY 678 AAGGCCCGAGTGGAGCCCTGACGAGCAGCTGC-----CACCACTGGTGGACGAGTGC 731
D 213 GluGluThrValAspSerIleValHisHisThrSerPheLysLysMetLysGluAsnCys 232
QY 732 ---TGCAACGCTGAGCCCTGCC-----GTG 755
D 233 MetThrAsnTyrThrIleProThrGluLeuMetAspHisAsnValSerProPheMet 252
QY 756 GCGCGGGAAGATTGGCTGTGGAGGACATCTTCCCGCTCTCCATCAATGAGAGTTT 815
D 253 ArgLysGlyThrThrGlyAspTrpLysAsnThrPheThrValAlaGlnAsnGluArgPhe 272
QY 816 GACTTGGTGTTATAACAGAGATGGAAAGTGTGACCTCAGCTTT 860
D 273 AspAlaHisTyrAlaLysThrMetThrAspCysAspPheLysPhe 287
```

RESULT 25

ADE57135

ID ADE57135 standard; protein; 291 AA.

XX AC ADE57135;

XX DT 29-JAN-2004 (first entry)

XX DE Rat Protein P17988, SEQ ID NO 2995.

XX KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
XX KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX OS Rattus norvegicus.

XX PN WO2003016475-A2.

XX PD 27-FEB-2003.

XX PF 14-AUG-2002; 2002WO-US025765.

XX PR 14-AUG-2001; 2001US-0312147P.

XX PR 01-NOV-2001; 2001US-0346382P.

XX PR 26-NOV-2001; 2001US-0333347P.

XX PA (GEHO) GEN HOSPITAL CORP.

XX PA (FARB) BAYER AG.

XX PI Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

XX DR GENBANK; P17988.

XX PT New composition comprising two or more isolated polypeptides, useful for
XX PS preparing a medicament for treating pain in an animal.

XX PS Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat
XX or human polynucleotides or a polynucleotide which represents a fragment,
XX derivative or allelic variation of the nucleic acid sequence. Also
XX claimed are a vector comprising the novel polynucleotide, a host cell
XX comprising the vector, a method for identifying a nucleotide sequence
XX which is differentially regulated in an animal subjected to pain and a
XX kit to perform the method, an array, a method for identifying an agent
XX that increases or decreases the expression of the polynucleotide sequence
XX that is differentially expressed in neuronal tissue of a first animal
XX subjected to pain, a method for identifying a compound which regulates
XX the expression of a polynucleotide sequence which is differentially
XX expressed in an animal subjected to pain, a method for identifying a
XX compound that regulates the activity of one or more of the
XX polynucleotides, a method for producing a pharmaceutical composition, a
XX method for identifying a compound or small molecule that regulates the
XX activity in an animal of one or more of the polypeptides given in the
XX specification, a method for identifying a compound useful in treating
XX pain and a pharmaceutical composition comprising the one or more
XX polypeptides or their antibodies. The polynucleotide or the compound that
XX modulates its activity is useful for preparing a medicament for treating
XX pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
XX injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
XX therapy). The sequence presented is a rat protein (shown in Table 2 of
XX the specification) which is differentially expressed during pain. Note:
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic form directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 291 AA;

Alignment Scores:

Pred. No.: 5.21e-34 Length: 291

Score: 458.00 Matches: 96

Percent Similarity: 55.6% Conservative: 57

Best Local Similarity: 34.9% Mismatches: 104

Query Match:	10.4%	Indels:	18
DB:	7	Gaps:	4
US-10-768-158-1 (1-2419) x ADE57135 (1-291)			
QY	90	GGCGTGGCGCTGCGCCCTTCTGCGCGGGAAGATGGAGGATCGCAACTTCCCGTG	149
DB	13	:	
QY	150	CGGCCAGCAGCTGTGTGATCGTACCTACCCCAAGTCCGGCACCAGCTTCTGCGAGGAG	209
DB	33	:	
QY	210	GTGTCTACTTGTGAGCAGGGCGCTGACCCCGATGAGATCGGCTTGATGAACATCGAC	269
DB	53	:	
QY	270	GAGCAGCTCCCGCTCTGAGTACCCACAGCG-----GGCCTGGACATCATC	317
DB	73	:	
QY	318	AAGAACTGACTCTCCCGCTCATCAAGAGCCACTGCGCTTCTGCGCTCT	377
DB	93	:	
QY	378	GACTCCCAATGAGAGTCCCAAGTCTATATGCTGCGCAACCCCAAGATCTGCTG	437
DB	113	:	
QY	438	GTGCTTATATAGTCCACCGCTCTCTGGGACCATGAGCTACCGAGGACCTTTC	497
DB	133	:	
QY	498	GAATTCCTCCGAGGTTTATGAATGAATGCTGCGCTACGGCTCTGTTTGAGCAGCTG	557
DB	153	:	
QY	558	CAGAGTCTGGGAGCACCGCATCGATCGAACGTGCTTTTCTCAAGATATGAAGCATG	617
DB	173	:	
QY	618	CATCGGACCTGGTGACATGCTGGAGCAGCTGCGCAGATTCCTGGGGTGCTCTGTGAC	677
DB	193	:	
QY	678	AAGCCCGAGCTGGAAGCCCTGACGAGCAGTCG-----CACCACTGGTGGACAGTGC	731
DB	213	:	
QY	732	---TGCAACGCTGAGCCCTGCGCC-----GTG	755
DB	233	MetThrAsnTyrThrThrIleProThrGluMetAspHisAsnValSerProPheMet	252
QY	756	GGCCGGGAGAGTGGCTGTGAGGACATCTCACCGTCTCCATGATGAGAGTTT	815
DB	253	ArgLysGlyThrThrGlyAspTrpLysAsnThrPheThrValAlaGlnAsnGluArgPhe	272
QY	816	GACTTGGTGATAACAGAAATGGAAAGTGTGACCTCACGTTT	860
DB	273	AspAlaHisTyrAlaLysThrMetThrAspCysaspPheLysPhe	287
RESULT 26			
ADE57147			
ID	ADE57147	standard; protein; 291 AA.	
XX	AC	ADE57147;	
XX	DT	29-JAN-2004 (first entry)	
XX	XX	Rat Protein P17988, SEQ ID NO 3007.	
XX	XX	Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;	
KW	XX	chronic constriction injury; CCI; spared nerve injury; SNI; Chung.	

XX	Rattus norvegicus.		
OS	WO2003016475-A2.		
PN	27-FEB-2003.		
PD	14-AUG-2002; 2002WO-US025765.		
PF	14-AUG-2001; 2001US-0312147P.		
XX	01-NOV-2001; 2001US-0346382P.		
XX	26-NOV-2001; 2001US-0333347P.		
XX	(GEHO) GEN HOSPITAL CORP.		
PA	(FARB) BAYER AG.		
XX	Woolf C, D'urso D, Befort K, Costigan M;		
PI	WPI; 2003-268312/26.		
XX	GENBANK; P17988.		
DR	New composition comprising two or more isolated polypeptides, useful for		
XX	preparing a medicament for treating pain in an animal.		
PT	Claim 1; Page; 1017pp; English.		
XX	The invention discloses a composition comprising two or more isolated rat		
CC	or human polynucleotides or a polynucleotide which represents a fragment,		
CC	derivative or allelic variation of the nucleic acid sequence. Also		
CC	claimed are a vector comprising the novel polynucleotide, a host cell		
CC	comprising the vector, a method for identifying a nucleotide sequence		
CC	which is differentially regulated in an animal subjected to pain and a		
CC	kit to perform the method, an array, a method for identifying an agent		
CC	that increases or decreases the expression of the polynucleotide sequence		
CC	that is differentially expressed in neuronal tissue of a first animal		
CC	subjected to pain, a method for identifying a compound which regulates		
CC	the expression of a polynucleotide sequence which is differentially		
CC	expressed in an animal subjected to pain, a method for identifying a		
CC	compound that regulates the activity of one or more of the		
CC	polynucleotides, a method for producing a pharmaceutical composition, a		
CC	method for identifying a compound or small molecule that regulates the		
CC	activity in an animal of one or more of the polypeptides given in the		
CC	specification, a method for identifying a compound useful in treating		
CC	pain and a pharmaceutical composition comprising the one or more		
CC	polypeptides or their antibodies. The polynucleotide or the compound that		
CC	modulates its activity is useful for preparing a medicament for treating		
CC	pain (e.g. spinal segmental nerve injury (Chung), chronic constriction		
CC	injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene		
CC	therapy). The sequence presented is a rat protein (shown in Table 2 of		
CC	the specification) which is differentially expressed during pain. Note:		
CC	The sequence data for this patent did not form part of the printed		
CC	specification, but was obtained in electronic form directly from WIPO at		
CC	ftp.wipo.int/pub/published_pct_sequences.		
XX	Sequence 291 AA;		
SQ			
Alignment Scores:			
Pred. No.:	5.21e-34	Length:	291
Score:	458.00	Matches:	96
Percent Similarity:	55.6%	Conservative:	57
Best Local Similarity:	34.9%	Mismatches:	104
Query Match:	10.4%	Indels:	18
DB:	7	Gaps:	4
US-10-768-158-1 (1-2419) x ADE57147 (1-291)			
QY	90	GGGTGGCGTGGCGCCCTTCTGCGCGGGAAGATGGAGGATCGCAACTTCCCGTG	149
DB	13	:	
QY	150	CGGCCAGCAGCTGTGTGATCGTACCTACCCCAAGTCCGGCACCAGCTTCTGCGAGGAG	209
DB	33	:	


```

Qy 210 GTGCTCTACTTGGTGAGCCAGGGCGCTGACCCCGATGAGATCGCGTTGATGAACATCGAC 269
Db 53 ILeuAspMetIleTyrGlnGlyGlyLeuGluLysCysGlyAlaProIleTyr 72
Qy 270 GAGCAGCTCCCGGTCCTGGAGTACCACAGCCG-----GGCCTGGACATCATC 317
Db 73 AlaArgValProPheLeuGluPheLysCysProGlyValProSerGlyLeuGluThrLeu 92
Qy 318 AAGGAACATGACTCTCCCGGCTCATCAAGAGCCACCTGCGCTTCTCCCTCT 377
Db 93 GluGluThrProAlaProArgLeuLeuLysThrHisLeuProLeuSerLeuLeuProGln 112
Qy 378 GACTCCCAATGAGAGATCCCAAGGTCTATATATGCTCGCAACCCCAAGGATCTGGTG 437
Db 113 SerLeuLeuAspGlnLysValLysValIleTyrIleAlaArgAsnAlaLysAspValVal 132
Qy 438 GTGCTTATTATAGTTCCACCGCTCTCTCGGACCATGAGCTACCGAGGACCTTTCAA 497
Db 133 ValSerTyrTyrAsnPheTyrAsnMetAlaLysLeuHisProAspProGlyThrTrpAsp 152
Qy 498 GAATTTCTCCGAGGAGTTTATGAATGATAACTGGGCTACGGCTCTCTGGTTGAGCAGTG 557
Db 153 SerPheLeuGluAsnPheMetAspGlyGluValSerTyrGlySerTyrTrpTyrGlnHisVal 172
Qy 558 CAGGAGTTCTGGAGCACCGCATGACTCGAAGCTGCTTTTCTCAAGTATGAAGACATG 617
Db 173 LysGluThrTrpGluLeuArgHisThrHisProValLeuTyrLeuPheTyrGluAspIle 192
Qy 618 CATCGGACCTGGTGACGATGCTGGAGCAGCTGGCCAGATTCCTGGGGGTCTCTGTGAC 677
Db 193 LysGluAsnProLysArgGluIleLysLysIleLeuGluPheLeuGlyArgSerLeuPro 212
Qy 678 AAGCCCAAGCTGGAAGCCCTGACGAGCATGCTC-----CACAGCTGGTGGACAGTGC 731
Db 213 GluGluThrValAspSerIleValHisHisThrSerPheLysLysMetLysGluAsnCys 232
Qy 732 ---TGCAACGCTGAGCCCTCGCC-----GTG 755
Db 233 MetThrAsnTyrThrThrIleProThrGluIleMetAspHisAsnValSerProPheMet 252
Qy 756 GCGCGGGAAGATTGGGCTGTGGAGGACATCTTCACCGTCTCATGAATGAGAAGTTT 815
Db 253 ArgLysGlyThrThrGlyAspTrpLysAsnThrPheThrValAlaGlnAsnGluArgPhe 272
Qy 816 GACTTGGTGTTAACAAGAGATGGGAAAGTGTGACCTCACCTTT 860
Db 273 AspAlaHisTyrAlaLysThrMetThrAspCysAspPheLysPhe 287
RESULT 27
AD57143
ID AD57143 standard; protein; 291 AA.
XX AC AD57143;
XX DT 29-JAN-2004 (first entry)
XX DE Rat Protein P17988, SEQ ID NO 3003.
XX KW Rat; pain; neuronal tissue; Gene therapy; spinal segmental nerve injury;
XX KW Chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX OS Rattus norvegicus.
XX FN WO2003016475-A2.
XX PD 27-FEB-2003.
XX PF 14-AUG-2002; 2002WO-US025765.
XX PR 14-AUG-2001; 2001US-0312147P.
XX PR 01-NOV-2001; 2001US-0346382P.
XX PR 26-NOV-2001; 2001US-0333347P.

```

```

XX (GEHO ) GEN HOSPITAL CORP.
XX (FARB ) BAYER AG.
XX Woolf C, D'urso D, Befort K, Costigan M;
XX MPI: 2003-268312/26.
XX GENBANK: P17988.
XX New composition comprising two or more isolated polypeptides, useful for
XX preparing a medicament for treating pain in an animal.
XX Claim 1; Page: 1017pp; English.
XX The invention discloses a composition comprising two or more isolated rat
XX or human polynucleotides or a polynucleotide which represents a fragment,
XX derivative or allelic variation of the nucleic acid sequence. Also
XX claimed are a vector comprising the novel polynucleotide, a host cell
XX comprising the vector, a method for identifying a nucleotide sequence
XX which is differentially regulated in an animal subjected to pain and a
XX kit to perform the method, an array, a method for identifying an agent
XX that increases or decreases the expression of the polynucleotide sequence
XX that is differentially expressed in neuronal tissue of a first animal
XX subjected to pain, a method for identifying a compound which regulates
XX the expression of a polynucleotide sequence which is differentially
XX expressed in an animal subjected to pain, a method for identifying a
XX compound that regulates the activity of one or more of the
XX polynucleotides, a method for producing a pharmaceutical composition, a
XX method for identifying a compound or small molecule that regulates the
XX activity in an animal of one or more of the polypeptides given in the
XX specification, a method for identifying a compound useful in treating
XX pain and a pharmaceutical composition comprising the one or more
XX polypeptides or their antibodies. The polynucleotide or the compound that
XX modulates its activity is useful for preparing a medicament for treating
XX pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
XX injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. Gene
XX therapy). The sequence presented is a rat protein (shown in Table 2 of
XX the specification) which is differentially expressed during pain. Note:
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic form directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 291 AA;
Alignment Scores:
Pred. No.: 5.21e-34 Length: 291
Score: 458.00 Matches: 96
Percent Similarity: 55.6% Conservative: 57
Best Local Similarity: 34.9% Mismatches: 104
Query Match: 10.4% Indels: 18
DB: 7 Gaps: 4
US-10-768-158-1 (1-2419) x AD57143 (1-291)
Qy 90 GCGTGGCGTCCGCGCTTCTGCGCGGGAAGATGGAGAGATCGCACTTCCCGTG 149
Db 13 GlyIleProLeuIleLysTyrPheAlaGluThrIleGlyProLeuGlnAsnPheThrAla 32
Qy 150 CGGCCCCAGCAGCTGTGGATCGTCACCTACCCCAAGTCCGCGACACAGCTTCTCGAGAG 209
Db 33 TrpProAspLeuLeuIleSerThrTyrProLysSerGlyThrThrTrpMetSerGlu 52
Qy 210 GTGCTCTACTTGGTGAGCCAGGGCGCTGACCCCGATGAGATCGCGTTGATGAACATCGAC 269
Db 53 ILeuAspMetIleTyrGlnGlyGlyLeuGluLysCysGlyAlaProIleTyr 72
Qy 270 GAGCAGCTCCCGGTCCTGGAGTACCACAGCCG-----GGCCTGGACATCATC 317
Db 73 AlaArgValProPheLeuGluPheLysCysProGlyValProSerGlyLeuGluThrLeu 92
Qy 318 AAGGAACATGACTCTCCCGGCTCATCAAGAGCCACCTGCGCTTCTCCCTCT 377
Db 93 GluGluThrProAlaProArgLeuLeuLysThrHisLeuProLeuSerLeuLeuProGln 112

```


Db 237 MetThrAsnTyrThrThrValArgGluPheMetAspHisSerIleSer----- 253
 Qy 747 CTGCCC---GTGGCCGGGGAAGAGTGGCGTGTGGGAAGGACATCTTCACCGTCTCCATG 803
 Db 254 ---ProPheMetArgLysGlyMetAlaGlyAspTrpLysThrThrPheThrValAlaGln 272
 Qy 804 AATGAGAGTTTGACTTGGTGTATATAACAGAGATGGGAAAGTGTGACCTCACGTTT 860
 Db 273 AsnGluArgPheAspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe 291

RESULT 29
 ADP65305
 ID ADP65305 standard; protein; 295 AA.
 XX ADP65305;
 AC ADP65305;
 XX 12-AUG-2004 (first entry)
 DT Human sulfotransferase family, cytosolic, 1a.
 XX
 DE
 XX autoimmune disease; arthritis; gene expression analysis;
 KW rheumatoid arthritis; collagen-induced; immunosuppressive; antirheumatic;
 KW antiarthritic; osteopathic; antigout; antiinflammatory; dermatological;
 KW immunomodulatory; lupus; ankylosing spondylitis; Fibrositis;
 KW fibromyalgia; osteoarthritis; gout; juvenile rheumatoid arthritis;
 KW immune; human.
 XX
 OS Homo sapiens.
 XX
 XX WO2003072827-A1.
 PN
 XX
 XX 04-SEP-2003.
 PD
 XX
 XX 31-OCT-2002; 2002MO-US035433.
 PF
 XX
 XX 31-OCT-2001; 2001US-0336220P.
 PR
 XX
 XX (CHIL-) CHILDREN'S HOSPITAL MEDICAL CENT.
 PA
 XX
 XX Hirsch R, Thorton SL;
 PI
 XX
 XX WPI; 2003-712740/67.
 DR
 XX
 XX GENBANK; NP_001045.
 DR
 XX
 XX Diagnosing and analyzing autoimmune disease using gene expression
 PT profiles and microarray technology, useful for diagnosing and treating
 FT rheumatoid arthritis, lupus, fibrositis, osteoarthritis, fibromyalgia and
 PT gout.
 PT
 XX
 XX Disclosure; Page; 56pp; English.
 PS
 XX
 XX The invention relates to a novel method for diagnosing and analysing
 CC autoimmune disease or arthritides. The method comprises obtaining a
 CC patient sample containing mRNA, analysing gene expression using the mRNA
 CC that results in a gene expression signature of the mRNA, and using that
 CC gene expression signature to diagnose or analyse the autoimmune disease
 CC or arthritides in the patient, where gene expression of at least 60% of
 CC the genes correlates with that of the gene signature. The invention
 CC further comprises: a treatment of rheumatoid arthritis; identification of
 CC genes for targeting in the treatment of rheumatoid arthritis in a mammal
 CC other than a mouse; diagnosis of rheumatoid arthritis in a mammal; an
 CC array or gene chip, specific for rheumatoid arthritis; diagnosis or
 CC analyses of autoimmune disease or rheumatoid arthritis; screening the
 CC efficacy of a candidate drug in vitro for the treatment of collagen-
 CC induced arthritis; and reducing the symptoms associated with collagen-
 CC induced arthritis. The compositions of the invention have the following
 CC activities: immunosuppressive, antirheumatic, antiarthritic, osteopathic,
 CC antigout, antiinflammatory, dermatological, and immunomodulatory. The
 CC methods and compositions of the present invention are useful for
 CC diagnosing and treating autoimmune disease or arthritides, such as
 CC rheumatoid arthritis, lupus, ankylosing spondylitis, fibrositis,
 CC fibromyalgia, osteoarthritis, gout, juvenile rheumatoid arthritis, and an
 CC immune disease caused by an infectious agent. This sequence represents a

CC protein sequence relating to the genes used in the analysis and treatment
 CC of autoimmune diseases or arthritides. Note: This sequence is not shown
 CC in the specification. It has been supplied in an electronic format from
 CC WIPO.
 XX
 SQ Sequence 295 AA;

Alignment Scores:
 Pred. No.: 8.14e-34 Length: 295
 Score: 456.00 Matches: 94
 Percent Similarity: 54.5% Conservative: 58
 Best Local Similarity: 33.7% Mismatches: 101
 Query Match: 10.4% Indels: 26
 DB: 7 Gaps: 4

US-10-768-158-1 (1-2419) x ADP65305 (1-295)

Qy 90 GGGGTGCGGTGCGCGCCCTTCTGCGCGGAAGATGGAGGATGCGCAACTTCCCGTG 149
 Db 17 GlyValProLeuIleLysTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla 36
 Qy 150 CGGCCAGCAGCGTGTGGATCGTCACCTACCCCAAGTCCGCCACCGCTGCTGCAGAG 209
 Db 37 ArgProAspAspLeuLeuIleSerThrTyrProLysSerGlyThrThrTrpValSerGln 56
 Qy 210 GTGCTCTACTTGGTGAGCGCGCTGACCCCGATGAGATCGGCTTGTGATGAACATCGAC 269
 Db 57 IleLeuAspMetIleTyrGlnGlyAspLeuGluLysCysHisArgAlaProIlePhe 76
 Qy 270 GAGCAGCTCCCGGTCTCTGGAGTACCCACAGCCG-----GGCTCGGACATCATC 317
 Db 77 MetArgValProPheLeuGluPheLysValProGlyIleProSerGlyMetGluThrLeu 96
 Qy 318 AAGAACTGACCTCTCCCGCTCATCAAGAGCCACCTGCGCTACCGCTTCTGCGCTCT 377
 Db 97 LysAsnThrProAlaProArgLeuLeuLysThrHisLeuProLeuAlaLeuLeuProGln 116
 Qy 378 GACCTCCACAAATGAGACTCCAAGTCTATATGCTCGCAACCCCAAGGATCTCGTG 437
 Db 117 ThrLeuLeuAspGlnLysValLysValValTyrValAlaArgAsnAlaLysAspValAla 136
 Qy 438 GTGTCTTATTATCAGTTCCACCGCTCTCTCGGACCATGAGTACCAGGACACCTTTCAA 497
 Db 137 ValSerTyrTyrHisPheTyrHisMetAlaLysValTyrProHisProGlyThrTrpGlu 156
 Qy 498 GAATTCGCGGAGGTTTATGAATGATAAGCTGGCTACGGCTCTGTTGACACATG 557
 Db 157 SerPheLeuGluLysPheMetAlaGlyValSerTyrGlySerTyrGlySerTyrGlnHisVal 176
 Qy 558 CAGGAGTCTGGGAGACCGCATGGACTCGAAGCTGCTTTTCTCAAGTAGTGAAGACATG 617
 Db 177 GlnGluTrpTrpGluLeuSerArgThrHisProValLeuTyrLeuPheTyrGluAspMet 196
 Qy 618 CATCGGACCTGTTGACGATGTTGAGACAGCTGGTGGACAGATTCCTGGGGGTGCTCTGAC 677
 Db 197 LysGluAsnProLysArgGluIleGlnLysIleLeuGluPheValGlyArgSerLeuPro 216
 Qy 678 AAGCCCCAGCTGGGAAGCCCTGACGAGCATGCG----- 710
 Db 217 GluGluThrValAspLeuMetValGluHisThrSerPheLysGluMetLysLysAsnPro 236
 Qy 711 -----CACAGCTGGTGGACCAAGTGTGCTGCAACGCTGAGGCC 746
 Db 237 MetThrAsnTyrThrThrValArgGluPheMetAspHisSerIleSer----- 253
 Qy 747 CTGCCC---GTGGCCGGGGAAGAGTGGCGTGTGGGAAGGACATCTTCACCGTCTCCATG 803
 Db 254 ---ProPheMetArgLysGlyMetAlaGlyAspTrpLysThrThrPheThrValAlaGln 272
 Qy 804 AATGAGAGTTTGACTTGGTGTATATAACAGAGATGGGAAAGTGTGACCTCACGTTT 860
 Db 273 AsnGluArgPheAspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe 291

RESULT 30
ADP24022
ID ADP24022 standard; protein; 295 AA.

XX AC ADP24022;

XX 18-NOV-2004 (first entry)

XX PRO polypeptide SEQ ID NO:1200.

XX PRO; antiinflammatory; antiarthritic; antirheumatic; immunosuppressive;
KW osteopathic; antidiabetic; dermatological; antipsoriatic; antiallergic;
KW antiaethmatic; hepatotropic; respiratory; gene therapy; immune system.

XX Unidentified.

XX WO2004041170-A2.

XX 21-MAY-2004.

XX 30-OCT-2003; 2003WO-US034312.

XX 01-NOV-2002; 2002US-0423394P.

XX (GETH) GENENTECH INC.

PI Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI;
PI Wu TD;

XX WPI; 2004-419628/39.

XX N-PSDB; ADP24021.

XX New PRO polypeptides and polynucleotides, useful for treating e.g.
PT erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated
PT renal disease, or demyelinating diseases of the central or peripheral
PT nervous system.

XX Claim 7; SEQ ID NO 1200; 2940pp; English.

XX The invention relates to a novel isolated nucleic acid and the PRO
CC polypeptide encoded by it. A protein of the invention has
CC antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,
CC osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,
CC antiaethmatic, hepatotropic, and respiratory activity. A polynucleotide
CC of the invention may have a use in gene therapy. The PRO polypeptide, its
CC agonist, antagonist, or antibody that specifically binds to the
CC polypeptide is useful for treating an immune related disorder such as
CC systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,
CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an
CC idiopathic inflammatory myopathy, Sjogren's syndrome, systemic
CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune
CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal
CC disease, a demyelinating disease of the central or peripheral nervous
CC system, idiopathic demyelinating polynuropathy, Guillain-Barre syndrome,
CC a chronic inflammatory demyelinating polynuropathy, a hepatobiliary
CC disease, infectious or autoimmune chronic active hepatitis, primary
CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,
CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's
CC disease, an autoimmune or immune-mediated skin disease, a bullous skin
CC disease, erythema multiforme, contact dermatitis, psoriasis, an allergic
CC disease, asthma, allergic rhinitis, atopic dermatitis, food
CC hypersensitivity, urticaria, an immunologic disease of the lung,
CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity
CC pneumonitis, a transplantation associated disease, graft rejection or
CC graft-versus-host disease. The present sequence represents a PRO protein
CC of the invention.

XX Sequence 295 AA;

Alignment Scores:
Pred. No.: 8.14e-34 Length: 295
Score: 456.00 Matches: 94
Percent Similarity: 54.5% Conservative: 58

Best Local Similarity: 33.7% Mismatches: 101
Query Match: 10.4% Indels: 26
DB: 8 Gaps: 4
US-10-768-158-1 (1-2419) x ADP24022 (1-295)

QY 90 GCGTGGCGGCTGCGCCCTTCTGCGCGGGAAGATGGAGGAGATCGCAACTTCCCGTG 149
DB 17 GlyValProLeuIleLysTyPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla 36
QY 150 CGGCCCGACGACGTGGATCGTACCTACCCCAAGTCCGCCACCGACTTCTGCAGGAG 209
DB 37 ArgProAspLeuLeuIleSerThrTyProLysSerGlyThrTrpValSerGln 56
QY 210 GTGCTCTACTTGGTGAGCCAGCGGCTGACCCCGATGAGATCGGCTTGTATGAACATCGAC 269
DB 57 IleLeuAspMetIleTyTGlnGlyAspLeuGlyLysCysHisArgAlaProIlePhe 76
QY 270 GAGCAGCTCCCGTCTGGAGTACCCACGCG-----GGCTGGACATCATC 317
DB 77 MetArgValProPheLeuGluPheLysValProGlyIleProSerGlyMetGluThrLeu 96
QY 318 AAGCAACTGACCTCTCCCGCTCATCAAGAGCCACCTGCCCTACCGCTTCTGCCCTCT 377
DB 97 LysAsnThrProAlaProArgLeuLeuLysHisLeuProLeuAlaLeuLeuProGln 116
QY 378 GACCTCCACAATGGAGACTCCCAAGTCTATATGGCTCGCAACCCCAAGGATCTGGTG 437
DB 117 ThrLeuLeuAspGlnLysValValTyValAlaArgAsnAlaLysAspValAla 136
QY 438 GTGCTTATTATCAGTTCCACCGCTCTCTCGGGACCATGAGTACCGAGGACCTTTCAA 497
DB 137 ValSerTyTyHisPheTyHisMetAlaLysValTyProHisProGlyThrTrpGlu 156
QY 498 GAATTCCTCCCGAGGTTTATGAATAGTACGGCTACGGCTCTGTTGTGGACACGTG 557
DB 157 SerPheLeuGluLysPheMetAlaGlyLysValSerTyGlySerTrpTyGlnHisVal 176
QY 558 CAGGAGTTCTGGGAGCACCGCATGGACTCGAAGCTGCTTTTCTCAAGTATGAAGACATG 617
DB 177 GlnGluTrpTrpGluLeuSerArgThrHisProValLeuTyLeuPheTyGluAspMet 196
QY 618 CATCGGACCTGGTGACGATGTGGAGCAGTGGCCAGATTCCTGGGGGTCTCTGTGAC 677
DB 197 LysGluAsnProLysArgGluIleGlnLysIleLeuGluPheValGlyArgSerLeuPro 216
QY 678 AAGSCCCAGCTGGAAGCCCTGACGAGCAGCTGC-----710
DB 217 GluGluThrValAspLeuMetValGluHisThrSerPheLysGluMetLysLysAsnPro 236
QY 711 -----CACGAGCTGGTGGACCAGCATGCTGCTGCAACGCTGAGGCC 746
DB 237 MetThrAsnTyThrThrValArgGluPheMetAspHisSerIleSer----- 253
QY 747 CTGCCC---GTGGCCCGGGAAGTGGGTGGGAGGAGCATCTTCAACGCTCTCCATG 803
DB 254 ---ProPheMetArgLysGlyMetAlaGlyAspTrpLysThrThrPheThrValaGln 272
QY 804 AATGAGAAGTTTGACTTGGTGTTATAACAGAGATGGAAAGTGTGACCTCACCTTT 860
DB 273 AsnGluArgPheAspAlaAspTyAlaGluLysMetAlaGlyCysSerLeuSerPhe 291

RESULT 31
ADW71800
ID ADW71800 standard; protein; 295 AA.
XX AC ADW71800;
XX DT 07-APR-2005 (first entry)
XX DE Pig phenol sulfotransferase protein, SULT1A1, SEQ ID 1.
XX KW Selectable marker; screening; animal breeding; polymorphism;

Db 157 SerPheLeuGluAspPheMetAlaGlyGluValSerTyrGlySerTrpTyrGlnHisVal 176
QY 558 CAGGAGTCTGGGAGCACCAGCATGCGAACTGCTTTTCTCAAGTATGAAGACATG 617
Db 177 GlnGluTrpTrpGluLeuArgHisThrHisProValLeuTyrLeuPheTyrGluAspMet 196
QY 618 CATCGGACCTGGTGACGATGGTGAGCAGCTGGCCAGATCTCTGGGGGTGCTGTGAC 677
Db 197 LysGluAsnProLysArgGluIleGlnLysIleLeuGluPheValGlyArgSerLeuPro 216
QY 678 AAGGCCAGCTGGAAGCCCTGACGGGACCTGCG-----CACACGCTGGTGGACCAG 728
Db 217 GluGluTrpValGluAspIleValGlnHisThrSerPheGlnGluMetLysAsnAsnAla 236
QY 729 TGCTGCAACGCTGAGGCCCTGCCCC-----GTG 755
Db 237 MetThrAsnTyrArgThrLeuProSerAspLeuLeuAspHisSerIleSerAlaPheMet 256
QY 756 GCGCGGGGAAGAGTTGGCTGTGGAGGACATCTTACCGCTCCCATGAATGAAGATTT 815
Db 257 ArgLysGlyIleThrGlyAspTrpLysSerThrPheThrValAlaGlnAsnGluArgPhe 276
QY 816 GACTTGGTGATAACAGAGATGGGAAGTGTGACCTCAGTTT 860
Db 277 GluAlaAspTyrAlaGluLysMetAlaGlyCysAsnLeuArgPhe 291
RESULT 35
ADD18776
ID ADD18776 standard; protein; 295 AA.
XX AC ADD18776;
XX DT 15-JAN-2004 (first entry)
XX DE Human disease related protein SeqID207.
XX KW human; disease state; cytostatic; antiinflammatory; ophthalmological;
KW antiarteriosclerotic; vulnery; gene therapy;
KW hypoxia-regulated condition; tumorigenesis; angiogenesis; apoptosis;
KW inflammation; erythropoiesis; glycolysis; gluconeogenesis;
KW glucose transportation; catecholamine synthesis; iron transport;
KW nitric oxide synthesis; cancer; ischaemic condition; reperfusion injury;
KW retinopathy; neonatal stress; pre-eclampsia; atherosclerosis;
KW inflammatory condition; wound healing.
XX OS Homo sapiens.
XX PN WO2003018621-A2.
XX PD 06-MAR-2003.
XX PF 23-AUG-2002; 2002WO-GB003892.
XX PR 23-AUG-2001; 2001GB-00020558.
XX PR 05-OCT-2001; 2001GB-00024037.
XX PA (OXFO-) OXFORD BIOMEDICA UK LTD.
XX PI Kingman SM, White J, Ward NR, Harris RA, Naylor S, Mundy CR;
XX DR N-PSDB; ADD18777.
XX DR 2003-290046/28.
XX PT New substantially purified polypeptide, useful for diagnosing or treating
XX PT a hypoxia-regulated condition, such as cancer, ischemia, reperfusion
XX PT injury, retinopathy, pre-eclampsia, atherosclerosis, inflammation, or
XX PT wound healing.
XX PS Claim 25; SEQ ID NO 207; 424pp; English.
XX CC This invention relates to novel human genes and gene product which are
XX CC implicated in certain disease states. Compounds which modulate the

CC proteins of the invention may have cytostatic, antiinflammatory,
CC ophthalmological, antiarteriosclerotic or vulnery activities. The
CC sequences of the invention may be useful for gene therapy. The invention
CC may be useful for diagnosing or treating a hypoxia-regulated condition,
CC such as tumorigenesis, angiogenesis, apoptosis, inflammation,
CC erythropoiesis, or the biological response to hypoxia conditions
CC including processes such as glycolysis, gluconeogenesis, glucose
CC transportation, catecholamine synthesis, iron transport or nitric oxide
CC synthesis. The disease includes cancer, ischaemic conditions, reperfusion
CC injury, retinopathy, neonatal stress, pre-eclampsia, atherosclerosis,
CC inflammatory conditions or wound healing. The present sequence is that of
XX a disease related protein of the invention.
SQ Sequence 295 AA;

Alignment Scores: 2,44e-33 Length: 295
Pred. No.: 451.00 Matches: 93
Score: 54.5% Conservative: 57
Percent Similarity: 33.8% Mismatches: 107
Best Local Similarity: 10.2% Indels: 18
Query Match: 7 Gaps: 3
DB:

US-10-768-158-1 (1-2419) x ADD18776 (1-295)
QY 90 GCGGTGCGGTGCGGCCCTTCTGCCCGGGAAGATGAGGAGATCGCAACTTCCCGGTG 149
Db 17 GlyValProLeuIleLysTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla 36
QY 150 CGGCCAGCAGCTGTGGATCGTCACCTACCCCAAGTCCGGCAGCAGCTGTGTCAGGAG 209
Db 37 ArgProAspLeuLeuIleSerThrTyrProLysSerGlyThrTrpValSerGln 56
QY 210 GTGCTACTGTGTGAGCCAGCGCTGACCCCGATGATGATGCGTGTGATGAACATCGAC 269
Db 57 IleLeuAspMetIleTyrGlnGlyAspLeuGluLysCysHisArgAlaProIlePhe 76
QY 270 GAGCAGCTCCCGTCTGTGAGTACCCACAGCCG-----GCCCTGGACATCATC 317
Db 77 MetArgValProPheLeuGluPheLysAlaProGlyIleProSerGlyMetGluThrLeu 96
QY 318 AAGGAAGTACCTCTCCCGCTCATCAAGAGCAGCTGCCCTACCGCTTCTGCGCTCT 377
Db 97 LysAspThrProAlaProArgLeuLeuLysThrHisLeuProLeuAlaLeuLeuProGln 116
QY 378 GACCTCCACAATGGAGCTCCAAGTCACTATATGGCTCGCAACCCCAAGGATCTGGTG 437
Db 117 ThrLeuLeuAspGlnLysValLysValValTyrValAlaArgAsnAlaLysAspValAla 136
QY 438 GTGCTTTATTCAGTTTCCACCGCTCTCTGCGGACCATGAGCTACCGAGGCACCTTTCAA 497
Db 137 ValSerTyrTyrHisPheTyrHisMetAlaLysValHisProGluProGlyThrTrpAsp 156
QY 498 GAATTCGCGGAGGTTTATGAATGATAGTGGGTACGGCTCGCTTGTGAGCAGCTG 557
Db 157 SerPheLeuGluLysPheMetValGlyValSerTyrGlySerTrpTyrGlnHisVal 176
QY 558 CAGGAGTCTGGGAGCACCAGCATGGAAGTCTGCAAGCTGCTTTCTCAAGTATGAAGACATG 617
Db 177 GlnGluTrpTrpGluLeuSerArgThrHisProValLeuTyrLeuPheTyrGluAspMet 196
QY 618 CATCGGACCTGTGACGATGGTGAGCAGCTGGCCAGATCTCTGGGGGTGCTGTGAC 677
Db 197 LysGluAsnProLysArgGluIleGlnLysIleLeuGluPheValGlyHisSerLeuPro 216
QY 678 AAGGCCAGCTGGAAGCCCTGACGGGACCTGCG-----CACACGCTGGTGGACCAG 728
Db 217 GluGluThrValAspPheMetValGlnHisThrSerPheLysGluMetLysLysAsnPro 236
QY 729 TGCTGCAACGCTGAGGCCCTGCCCC-----GTG 755
Db 237 MetThrAsnTyrThrThrValProGlnGluPheMetAspHisSerIleSerProPheMet 256

QY 756 GGCGGGGAGAGTTGGCTGTCGAGGACATCTTACCGCTCCCATGAATGAGAGTTT 815
DB 257 ArgLysGlyMetAlaGlyAspTrpLysThrPheThrValAlaGlnAenGluArgPhe 276
QY 816 GACTTGGTGATATAACAGAGATGGGAAAGTGTGACCTCACGTTT 860
DB 277 AspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe 291
RESULT 36
ID ADE57149
AD ADE57149 standard; protein; 295 AA.
XX ADE57149;
AC ADE57149;
XX 29-JAN-2004 (first entry)
DT Human Protein P50225, SEQ ID NO 3009.
DE Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX Homo sapiens.
XX W02003016475-A2.
XX 27-FEB-2003.
XX 14-AUG-2002; 2002WO-US025765.
XX 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX (GHEO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX Woolf C, D'urso D, Befort K, Costigan M;
XX WPI; 2003-268312/26.
DR GENBANK; P50225.
XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX Claim 1; Page; 1017pp; English.
XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative of allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (SNI)), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 295 AA;
SQ
Alignment Scores: 2,44e-33 Length: 295
Pred. No.: 451.00 Matches: 93
Score: 54.5% Conservatives: 57
Percent Similarity: 33.8% Mismatches: 107
Best Local Similarity: 10.2% Indels: 18
Query Match: 7 Gaps: 3
DB:
US-10-768-158-1 (1-2419) x ADE57149 (1-295)
QY 90 GGCGTGGCTGCCGCCCTTCTCCCGGGAGAGATGGAGGAGATGCCAACTTCCCGGTG 149
DB 17 GlyValProLeuIleLysTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla 36
QY 150 CGGCCAGGAGCGTGTGGATCGTCACCTACCCCAAGTCCGGCACCAGCTTGTCTGACGAG 209
DB 37 ArgProAspLeuLeuIleSerThrTyrProLysSerGlyThrThrTrpValSerGln 56
QY 210 GTGGTCTACTTGTGAGCCAGCGCTGACCCCGATGATCGCTTGTGAACATCGAC 269
DB 57 IleLeuAspMetIleTyrGlnGlyGlyAspLeuGluLysCysHisArgAlaProIlePhe 76
QY 270 GAGCAGCTCCCGTCTCTGGAGTACCCACAGCCG-----GGCTGGACATCATC 317
DB 77 MetArgValProPheLeuGluPheLysAlaProGlyIleProSerGlyMetGluThrLeu 96
QY 318 AAGGAAGTACCTCTCCCGCTCATCAAGAGCCACCTGCCCTACCGCTTCTGCCCTCT 377
DB 97 LysAspThrProAlaProArgLeuLeuLysThrHisLeuProLeuAlaLeuProGln 116
QY 378 GACCTCCACAATGGAGACTCCAAAGGTCTATCTATATGCTCGCAACCCCAAGGATCTGGT 437
DB 117 ThrLeuLeuAspGlnLysValValValValValValValValValValValValVal 136
QY 438 GTGCTTTATTCAGTTTCCACCGCTCTCTGCGGACCATGAGCTACCGAGGCACCTTCAA 497
DB 137 ValSerTyrTyrHisPheTyrHisMetAlaLysValHisProGluProGlyThrTrpAsp 156
QY 498 GAATTCGCGGAGGTTTATGAATATAAGCTCGGCTACGGCTCTGGTTTGGACGAGTG 557
DB 157 SerPheLeuGluLysPheMetValGlyGluValSerTyrGlySerTyrGlnHisVal 176
QY 558 CAGGAGTCTGGGAGCACCAGCATGGACTCGAAGCTGCTTTTCTCAAGTATGAAGCATG 617
DB 177 GlnGluTrpTrpGluLeuSerArgThrHisProValLeuTyrLeuPheTyrGluAspMet 196
QY 618 CATCGGACCTGGTGAGATGGTGGAGCAGCTGGCCAGATTCCTGGGGGTGCTCTGTGAC 677
DB 197 LysGluAsnProLysArgGluIleGlnLysIleLeuGluPheValGlyArgSerLeuPro 216
QY 678 AAGGCCAGCTGGAAGCCCTGACGGAGCACTGC-----CACCAGCTGGTGGACAG 728
DB 217 GluGluThrValAspPheMetValGlnHisThrSerPheLysGluMetLysLysAsnPro 236
QY 729 TGCTCAACCTCGAGGCGCTGCC-----GTG 755
DB 237 MetThrAsnTyrThrThrValProGlnGluPheMetAspHisSerIleSerProPheMet 256
QY 756 GGCGGGGAGAGTTGGCTGTCGAGGACATCTTACCGCTCCCATGAATGAGAGTTT 815
DB 257 ArgLysGlyMetAlaGlyAspTrpLysThrPheThrValAlaGlnAenGluArgPhe 276
QY 816 GACTTGGTGATATAACAGAGATGGGAAAGTGTGACCTCACGTTT 860
DB 277 AspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe 291
RESULT 37
ADE57137
ID ADE57137 standard; protein; 295 AA.

XX	AD557137;	US-10-768-158-1 (1-2419) x AD557137 (1-295)	
XX	29-JAN-2004 (first entry)	90 GGCCTGGCGTGGCGGCTCTCTGCGGGGAAGATGGAGGAGATCGCAACTTCCCGGTG 149	
XX	Human Protein P50225, SEQ ID NO 2997.	17 GlyValProLeuIleLeuTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla 36	
XX	Human; pain; neuronal tissue; gene therapy;	150 CGGCCCGGACGACGTGTGGATCGTCACCTACCCCAAGTCCGGCAGCAGCTTCTCGCAGAG 209	
XX	spinal segmental nerve injury; chronic constriction injury; CCI;	37 ArgProAspLeuLeuIleSerThrTyrProLysSerGlyThrTrpLysValSerGln 56	
XX	spared nerve injury; SNI; Chung.	210 GTGCTCTACTTGGTGAGCCAGGCGCTGACCCCGATGAGATCGGCTTCATCAACATCCAC 269	
XX	Homo sapiens.	57 IleLeuAspMetIleTyrGlnGlyGlyAspLeuGlnLysCysHisArgAlaProIlePhe 76	
XX	WO2003016475-A2.	270 GAGCAGCTCCCGGTCTCTGGAGTACCCACAGCCG-----GGCCTGGACATCATC 317	
XX	27-FEB-2003.	77 MetArgValProPheLeuGluPheLysAlaProGlyIleProSerGlyMetGluThrLeu 96	
XX	14-AUG-2002; 2002WO-US025765.	318 AAGAACTGACCTCTCCCGCTCATCAAGAGCCACCTGCCTACCGCTTCTGCCCTCT 377	
XX	14-AUG-2001; 2001US-0312147P.	97 LysAspThrProAlaProArgLeuLeuLysThrHisLeuProLeuAlaLeuLeuProGln 116	
XX	01-NOV-2001; 2001US-0346382P.	378 GACCTCCACAATCGAGACTCCAGGTCTATATGCTCGCAACCCCAAGATCTGGTG 437	
XX	26-NOV-2001; 2001US-0333347P.	117 ThrLeuLeuAspGlnLysValLysValValTyrValAlaArgAsnAlaLysAspValAla 136	
XX	(GEHO) GEN HOSPITAL CORP.	438 GTGCTTTATTATCAGTTTCCACCGCTCTCTGCGGACCATGAGCTACCGAGGACCTTTCAA 497	
XX	(FARB) BAYER AG.	137 ValSerTyrThrHisPheTyrHisMetAlaLysValHisProGluProGlyThrTrpAsp 156	
XX	Woolf C, D'urso D, Befort K, Costigan M;	498 GAATTTCTCCGGAGTTTATGAATGATGAAGCTGGGCTACCGCTCCTGGTTTGACAGCTG 557	
XX	WPI; 2003-268312/26.	157 SerPheLeuGluLysPheMetValGlyGluValSerTyrGlySerTyrTyrGlnHisVal 176	
XX	GENBANK; P50225.	558 CAGGAGTTCTGGGAGCACCAGTCGACTCGAACGTCCTTTTCTCAAGTATGAGACATG 617	
XX	New composition comprising two or more isolated polypeptides, useful for	177 GlnGluTrpTrpGluLeuSerArgThrHisProValLeuTyrLeuPheTyrGluAspMet 196	
XX	preparing a medicament for treating pain in an animal.	618 CATCGGACCTGGTGACGATGGTGAGCAGCTGGCCAGATTCCTGGGGGTCTCCTGTGAC 677	
XX	Claim 1; Page; 1017pp; English.	197 LysGluAsnProLysArgGluIleGlnLysIleLeuGluPheValGlyArgSerLeuPro 216	
XX	The invention discloses a composition comprising two or more isolated rat	678 AAGCCCGCAGCTGGAAGCCCTGACGAGCAGCTGC-----CACCAGCTGGTGGACAG 728	
XX	or human polynucleotides or a polynucleotide which represents a fragment,	217 GluGluThrValAspPheMetValGlnHisThrSerPheLysGluMetLysLysAsnPro 236	
XX	derivative or allelic variation of the nucleic acid sequence. Also	729 TGCTGCAACGCTGAGGCCCTGCCCC-----GTG 755	
XX	claimed are a vector comprising the novel polynucleotide, a host cell	237 MetThrAsnTyrThrThrValProGlnGluPheMetAspHisSerIleSerProPheMet 256	
XX	comprising the vector, a method for identifying a nucleotide sequence	756 GGCCGGGAAGAGTTGGCTGTGGAAGGACATCTTACCGCTCTCCATGAATGAGAAGTTT 815	
XX	which is differentially regulated in an animal subjected to pain and a	257 ArgLysGlyMetAlaGlyAspTrpLysThrThrPheThrValAlaGlnAsnGluArgPhe 276	
XX	kit to perform the method, an array, a method for identifying an agent	816 GACTTGGTGTATAACAGAGATGGAAAGTGGAGCTGACCTCAGCTTT 860	
XX	that increases or decreases the expression of the polynucleotide sequence	277 AspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe 291	
XX	which is differentially expressed in neuronal tissue of a first animal	RESULT 38	
XX	the expression of a polynucleotide sequence which is differentially	AD557141	
XX	expressed in an animal subjected to pain, a method for identifying a	ID AD557141 standard; protein; 295 AA.	
XX	compound that regulates the activity of one or more of the	XX AC AD557141;	
XX	polynucleotides, a method for producing a pharmaceutical composition, a	XX DT 29-JAN-2004 (first entry)	
XX	method for identifying a compound or small molecule that regulates the	XX DE Human Protein P50225, SEQ ID NO 3001.	
XX	activity in an animal of one or more of the polypeptides given in the	XX KW Human; pain; neuronal tissue; gene therapy;	
XX	specification, a method for identifying a compound useful in treating	XX KW spinal segmental nerve injury; chronic constriction injury; CCI;	
XX	pain and a pharmaceutical composition comprising the one or more	XX KW spared nerve injury; SNI; Chung.	
XX	polypeptides or their antibodies. The polynucleotide or the compound that		
XX	modulates its activity is useful for preparing a medicament for treating		
XX	pain (e.g. spinal segmental nerve injury (Chung), chronic constriction		
XX	injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene		
XX	therapy). The sequence presented is a human protein (shown in Table 2 of		
XX	the specification) which is differentially expressed during pain. Note:		
XX	The sequence data for this patent did not form part of the printed		
XX	specification, but was obtained in electronic form directly from WIPO at		
XX	ftp.wipo.int/pub/published_pct_sequences.		
XX	Sequence 295 AA;		
XX	Alignment Scores:		
XX	Pred. No.:	2.44e-33	Length: 295
XX	Score:	451.00	Matches: 93
XX	Percent Similarity:	54.5%	Conservative: 57
XX	Best Local Similarity:	33.8%	Mismatches: 107
XX	Query Match:	10.2%	Indels: 18
XX	DB:	7	Gaps: 3

OS Homo sapiens.
 XX WO2003016475-A2.
 XX PD 27-FEB-2003.
 XX PF 14-AUG-2002; 2002WO-US025765.
 XX PR 14-AUG-2001; 2001US-0312147P.
 PR 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-0333347P.
 XX (GEO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.
 XX Woolf C, D'urso D, Befort K, Costigan M;
 FI WPI; 2003-268312/26.
 XX GENBANK; P50225.
 DR
 DR
 DR
 XX New composition comprising two or more isolated polypeptides, useful for
 PT preparing a medicament for treating pain in an animal.
 PT
 XX
 XX
 PS Claim 1; Page: 1017pp; English.
 XX
 CC The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a human protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 295 AA;

Alignment Scores:
 Pred. No.: 2,44e-33 Length: 295
 Score: 451.00 Matches: 93
 Percent Similarity: 54.5% Conservative: 57
 Best Local Similarity: 33.8% Mismatches: 107
 Query Match: 10.2% Indels: 18
 DB: 7 Gaps: 3

US-10-768-158-1 (1-2419) x ADE57141 (1-295)

QY 90 GCGGTGCGGCTCGCCCTCTCTCGCGCGGAAGATGGAGGAGATCGCAACTTCCCGGTG 149
 Db 17 GlyValProLeuIleLysTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnIla 36
 QY 150 CGGCCAGCGAGTGTGGATGCTACCTACCCCAAGTCGGGACCCAGCTTGTCTGCGAGAG 209
 Db 37 ArgProAspLeuLeuIleSerThrTyrProLysSerGlyThrThrTrpValSerGln 56

QY 210 GTGCTCTACTTGTGAGCGCGGTGACCCCGATGAGATCGGCTTGTATGAACATCGAC 269
 Db 57 IleLeuAspMetIleTyrGlnGlyGlyAspLeuGlnLysCysHisArgAlaProIlePhe 76
 QY 270 GAGCAGCTCCCGGTCTCTGGAGTAGTACCCACACCGC-----GCCCTGGACATCATC 317
 Db 77 MetArgValProPheLeuGluPheLysAlaProGlyIleProSerGlyMetGluThrLeu 96
 QY 318 AAGAACTGACCTCTCCCGCTCATCAAGAGCAGCTCCCTACCGCTTCTGCTCTCT 377
 Db 97 LysAspThrProAlaProArgLeuLeuLysThrHisLeuProLeuAlaLeuLeuProGln 116
 QY 378 GACCTCCACATGGAGACTCCAAGGTCACTATATGCTCGCAAGCCCAAGGATCTCGTG 437
 Db 117 ThrLeuLeuAspGlnLysValLysValValValValAlaArgAlaLysAspValAla 136
 QY 438 GTGTCTTATTATCATGTTTCCACCGCTCTCTCGGACCATGAGTACCGAGGACACCTTTCAA 497
 Db 137 ValSerTyrTyrHisPheTyrHisMetAlaLysValHisProGluProGlyThrTrpAsp 156
 QY 498 GAATCTCGCGAGGTATTATGAATGATAGCTGGGCTACGCTCTCGTGGTGTGACAGTG 557
 Db 157 SerPheLeuGluLysPheMetValGlyValSerTyrGlySerTyrTyrGlnHisVal 176
 QY 558 CAGGAGTCTGGGAGCACCGCATGCGACTCGAAACGTGCTTTTCTCAAGTATGAAGACATG 617
 Db 177 GlnGluTrpGluLeuSerArgThrHisProValLeuTyrLeuPheTyrGluAspMet 196
 QY 618 CATCGGACCTGCTGAGCATGTTGGAGCAGCTGGCCAGATTCCTGGGGGTGCTCTGTGAC 677
 Db 197 LysGluAsnProLysArgGluIleGlnLysIleLeuGluPheValGlyArgSerLeuPro 216
 QY 678 AAGCCCGAGCTGGAAGCCCTGACGGACCATGCG-----CACCAGCTGGTGGACAG 728
 Db 217 GluGluThrValAspPheMetValGlnHisThrSerPheLysGluMetLysLysAsnPro 236
 QY 729 TGCTGCAACGCTGAGGCCCTGCGCC-----GTG 755
 Db 237 MetThrAsnTyrThrThrValProGlnGluPheMetAspHisSerIleSerProPheMet 256
 QY 756 GGCCGGGAAGAGTGGGCTGTGGAGGACATCTTACCGCTCTCCATGAATGAGAAGTTT 815
 Db 257 ArgLysGlyMetAlaGlyAspTrpLysThrThrPheThrValAlaGlnAsnGluArgPhe 276
 QY 816 GACTTGTGTATATAACAGAGATGGGAAGTGTGACCTCAGCTT 860
 Db 277 AspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe 291
 RESULT 39
 ADE57145
 ID ADE57145 standard; protein; 295 AA.
 XX
 XX ADE57145;
 XX
 XX 29-JAN-2004 (first entry)
 XX
 DE Human Protein P50225, SEQ ID NO 3005.
 XX
 KW Human; pain; neuronal tissue; gene therapy;
 KW spinal segmental nerve injury; chronic constriction injury; CCI;
 KW spared nerve injury; SNI; Chung.
 XX
 OS Homo sapiens.
 XX
 XX WO2003016475-A2.
 XX PD 27-FEB-2003.
 XX PF 14-AUG-2002; 2002WO-US025765.
 XX PR 14-AUG-2001; 2001US-0312147P.
 PR 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-0333347P.

Qy	618	CATCGGACCTGGTGACCATGGTGTGAGCAGCTGGCCAGATTCTCTGGGGTGTCTCTGTGTGAC	677
Db	197	LysGluAsnProLysArgGluIleGlnLysIleLeuGluPheValGlyHisSerLeuPro	216
Qy	678	AAGGCCACAGCTGGGAAGCCTGACGGAGCACTGC-----CACCAGCTGGTGACACG	728
Db	217	GluGluThrValAspPheMetValGlnHisThrSerPheLysGluMetLysLysAsnPro	236
Qy	729	TGCTGCAACGCTGAGGCCCTGCC-----GTG 755	
Db	237	MetThrAsnTyrThrThrValProGlnGluPheMetAspHisSerIleSerProPheMet	256
Qy	756	GGCCGGGGAAGAGTGGCTGTGGGAAGGACATCTTCACCGCTCTCCATGAATGAGAAGTTT	815
Db	257	ArgLysGlyMetAlaGlyAspTrpLysThrThrPheThrValAlaGlnAsnGluArgPhe	276
Qy	816	GACTTGGTGTTAAACAGAAAGTGGAAAGTGTGACCTCACGCTT 860	
Db	277	AspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe	291
RESULT	44		
AD117050			
ID	AD117050	standard; protein; 302 AA.	
XX	AC	AD117050;	
XX	XX		
DT	15-APR-2004	(first entry)	
XX	XX		
DE	Human NOVX	protein homologue SeqID 586.	
XX	XX		
KW	human; NOVX;	cardiomyopathy; atherosclerosis; cancer; diabetes;	
KW	inflammation; autoimmune disorder;	allergy; blood disorder;	
KW	acquired immunodeficiency syndrome; AIDS;	obesity; aschma;	
KW	immunoglobulin (Ig) A nephropathy;	cirrhosis; arthritis;	
KW	Alzheimer's disease;	infection; str.	
XX	OS	Homo sapiens.	
XX	XX		
PN	WO200268649-A2.		
XX	XX		
PD	06-SEP-2002.		
XX	XX		
PF	31-JAN-2002;	2002WO-US002785.	
XX	XX		
PR	31-JAN-2001;	2001US-0265395P.	
PR	31-JAN-2001;	2001US-0265412P.	
PR	31-JAN-2001;	2001US-0265514P.	
PR	31-JAN-2001;	2001US-0265517P.	
PR	02-FEB-2001;	2001US-0266406P.	
PR	05-FEB-2001;	2001US-0266767P.	
PR	07-FEB-2001;	2001US-0266975P.	
PR	07-FEB-2001;	2001US-0267057P.	
PR	08-FEB-2001;	2001US-0267459P.	
PR	09-FEB-2001;	2001US-0267823P.	
PR	15-FEB-2001;	2001US-0268974P.	
PR	26-FEB-2001;	2001US-0271664P.	
PR	27-FEB-2001;	2001US-0271839P.	
PR	27-FEB-2001;	2001US-0271855P.	
PR	02-MAR-2001;	2001US-0272788P.	
PR	02-MAR-2001;	2001US-0273046P.	
PR	14-MAR-2001;	2001US-0275025P.	
PR	14-MAR-2001;	2001US-0275947P.	
PR	14-MAR-2001;	2001US-0275950P.	
PR	14-MAR-2001;	2001US-0275989P.	
PR	15-MAR-2001;	2001US-0276448P.	
PR	15-MAR-2001;	2001US-0276450P.	
PR	16-MAR-2001;	2001US-0276397P.	
PR	16-MAR-2001;	2001US-0276768P.	
PR	20-MAR-2001;	2001US-0278652P.	
PR	26-MAR-2001;	2001US-0278775P.	
PR	26-MAR-2001;	2001US-0278778P.	
PR	29-MAR-2001;	2001US-0279882P.	
PR	29-MAR-2001;	2001US-0279884P.	

Score:	449.00	Matches:	96
Percent Similarity:	55.8%	Conservative:	52
Best Local Similarity:	36.2%	Mismatches:	99
Query Match:	5	Indels:	18
DB:		Gaps:	3
US-10-768-158-1 (1-2419) x ADI17050 (1-302)			
Qy	126	GAGGAGATCGCAACTTCCCGGTGGGCCCGCCAGCGAGCGTGCGTGGATCGTCACCTACCCCAAG	185
Db	36	AspLysIleTrpAsnPheGlnAlaLysProAspAspLeuLeuIleSerThrTyProLys	55
Qy	186	TCCGGCACCACTTGTCTCAGAGGTGCTACTTGTGTGAGCCAGGGCGCTGACCCGAT	245
Db	56	AlaGlyThrThrTrpThrGlnGluIleValGluLeuIleGlnAsnGluGlyAspValGlu	75
Qy	246	GAGATCGCGCTTGATGAACATCGAGCAGCGCTCCCGGTCTCTGGAGTACCCACACCGC	302
Db	76	LysSerLysArgAlaProThrHisGlnArgPheProPheLeuGluMetLysIleProSer	95
Qy	303	-----GCCCTGGACATCATCAAGAACTGACCTCTCCCGCTCATCAAGACGAC	353
Db	96	LeuGlySerGlyLeuGluGlnAlaHisAlaMetProSerProArgIleLeuLysThrHis	115
Qy	354	CTGCCCTACCGCTTCTGCCCTCTGACCTCCACAATGGAGACTCCAGGCTCATCTATATG	413
Db	116	LeuProPheHisLeuLeuProProSerLeuLeuGluLysAsnCysLysIleIleTyVal	135
Qy	414	GCTCGCAACCCCAAGGATCTGGTGGTCTTATTATCATGTTTCCACCGCTCTCTCGGACC	473
Db	136	AlaArgAsnProLysAspAsnMetValSerTyTrpHisPheGlnArgMetAsnLysAla	155
Qy	474	ATGAGCTACCGAGGACCTTTCAAGATCTGCCGGAGGTTTATGAATGATAGCTGGGC	533
Db	156	LeuProAlaProGlyThrTrpGluGluTrpPheGluThrPheLeuAlaGlyLysValCys	175
Qy	534	TACGGCTCTGGTTGAGCAGCTGAGGAGTTCTGGGAGCACCGCATGGACTCGAAGCTG	593
Db	176	TrpGlySerTrpHisGluHisValLysGlyTrpTrpGluAlaLysAspLysHisArgIle	195
Qy	594	CTTTTCTTCAAGTATGAAGACATCGACCTGGACCTGTGTGAGGTGGAGCAGCTGGCC	653
Db	196	LeuTyLeuPheTyLeuGluAspMetLysLysAsnProLysHisGluIleGlnLysLeuAla	215
Qy	654	AGATCTCGGGGTCTCTGTGACAGCCCGACGCTGGAAGCCCTGACGGACATCGCAC	713
Db	216	GluPheIleGlyLysLysLeuAspLysValLeuAspLysIleValHisTyThrSer	235
Qy	714	CAGCTGGTG-----GACAGTGTGCAACGCTGAGGCCCTGCC-----	752
Db	236	PheAspValMetLysGlnAsnProMetAlaAsnTySerSerIleProAlaGluIleMet	255
Qy	753	-----GTGGGCCGGGAAGTGTGGGCTGTGGAAGACATCTTC	791
Db	256	AspHisSerIleSerProPheMetArgLysGlyAlaValGlyAspTrpLysLysHisPhe	275
Qy	792	ACCGTCTCCATGATGAGAATTTGACTTGTGTGTATATAACAGACATCGGAAAGTGTCAC	851
Db	276	ThrValAlaGlnAsnGluArgPheAspGluAspTyTyLysLysMetThrAspThrArg	295
Qy	852	CTCAGCTTGACTTT 866	
Db	296	LeuThrPheHisPhe 300	
RESULT 45			
ADF76915			
ID	ADF76915	standard; protein; 302 AA.	
XX			
AC	ADF76915;		
XX			
DT	26-FEB-2004	(first entry)	
XX			
DE	Novel human secreted and transmembrane protein seqID 590.		
XX	human; PRO; membrane bound protein; membrane bound receptor;		
KW	cell proliferation; cell migration; cell differentiation;		
KW	mitogenic factor; survival factor; cytotoxic factor;		
KW	differentiation factor; neuro peptide; hormone; cell receptor;		
KW	receptor-ligand interaction; cytostatic; chondrocyte; tumour.		
OS	Homo sapiens.		
XX			
FN	WO2003072035-A2.		
XX			
PD	04-SEP-2003.		
XX			
PF	21-FEB-2003; 2003WO-US005241.		
XX			
PR	22-FEB-2002; 2002US-0359461P.		
XX			
PA	(GETH) GENENTECH INC.		
XX			
PI	Bodary SC, Clark H, Hunte B, Jackman JK, Schoenfeld JR;		
PI	Williams PM, Wood WI, Wu TD;		
XX			
DR	WPI; 2003-721702/68.		
DR	N-PSDB; ADF76914.		
XX			
PT	New PRO polypeptides, useful for diagnosing and treating an immune		
PT	related disorder, e.g. systemic lupus erythematosus, rheumatoid		
PT	arthritis, osteoarthritis, juvenile chronic arthritis, thyroiditis or		
PT	diabetes mellitus.		
XX			
PS	Claim 10; SEQ ID NO 590; 918pp; English.		
XX			
CC	This invention relates to novel nucleic acids encoding human PRO secreted		
CC	and transmembrane proteins. Extracellular proteins play important roles		
CC	in the formation, differentiation and maintenance of multicellular		
CC	organisms. The fate of many individual cells (for example proliferation,		
CC	migration or differentiation) is typically governed by information		
CC	received from other cells and the immediate environment. The information		
CC	is often transmitted by secreted polypeptides (for example mitogenic		
CC	factors, survival factors, cytotoxic factors, differentiation factors,		
CC	neuropeptides or hormones) which are received and interpreted by diverse		
CC	cell receptors or membrane bound proteins. These membrane bound proteins		
CC	and receptors may be of use as pharmaceutical and diagnostic agents, such		
CC	as in the blocking of receptor-ligand interactions. The current invention		
CC	provides the amino acid sequences of novel human membrane bound receptors		
CC	and proteins, along with the cDNA sequences encoding them. The novel		
CC	proteins of the invention may have cytostatic activities through the		
CC	stimulation of chondrocytes. The nucleic acids of the invention may be		
CC	useful for the manufacture of a medicament for diagnosing or treating a		
CC	tumour in a mammal. In addition, they may be useful for measuring or		
CC	detecting the expression of a tumour associated gene. The present		
CC	sequence is the amino acid sequence of a human PRO protein of the		
CC	invention.		
XX			
SQ	Sequence 302 AA;		
Alignment Scores:			
Pred. No.:	3.82e-33	Length:	302
Score:	449.00	Matches:	96
Percent Similarity:	55.8%	Conservative:	52
Best Local Similarity:	36.2%	Mismatches:	99
Query Match:	10.2%	Indels:	18
DB:	7	Gaps:	3
US-10-768-158-1 (1-2419) x ADF76915 (1-302)			
Qy	126	GAGGAGATCGCAACTTCCCGGTGGGCCCGCCAGCGAGCGTGCGTGGATCGTCACCTACCCCAAG	185
Db	36	AspLysIleTrpAsnPheGlnAlaLysProAspAspLeuLeuIleSerThrTyProLys	55
Qy	186	TCCGGCACCACTTGTCTCAGAGGTGCTACTTGTGTGAGCCAGGGCGCTGACCCGAT	245
Db	56	AlaGlyThrThrTrpThrGlnGluIleValGluLeuIleGlnAsnGluGlyAspValGlu	75

```

Qy 246 GAGATCGGCTTGATGAACATCAGACGAGCAGCTCCGGTCTCGGAGTACCCACAGCCG--- 302
Db 246 GAGATCGGCTTGATGAACATCAGACGAGCAGCTCCGGTCTCGGAGTACCCACAGCCG--- 302
Qy 76 LysSerLysArgAlaProThrHisGlnArgPheProPheLeuGluMetLysIleProSer 95
Db 76 LysSerLysArgAlaProThrHisGlnArgPheProPheLeuGluMetLysIleProSer 95
Qy 303 -----GGCTGGACATCATCAGAACTGACCTCCCGCTCATCAAGAGCCAC 353
Db 303 -----GGCTGGACATCATCAGAACTGACCTCCCGCTCATCAAGAGCCAC 353
Qy 96 LeuGlySerGlyLeuGluGlnAlaHisAlaMetProSerProArgIleLeuLysThrHis 115
Db 96 LeuGlySerGlyLeuGluGlnAlaHisAlaMetProSerProArgIleLeuLysThrHis 115
Qy 354 CTGCCCTACCGCTTCTCCCTCTGACCTCCACAAATGAGAGCTCAAGAGTTCATATATG 413
Db 354 CTGCCCTACCGCTTCTCCCTCTGACCTCCACAAATGAGAGCTCAAGAGTTCATATATG 413
Qy 116 LeuProPheHisLeuLeuProSerLeuLeuGluLysAsnCysLysIleIleTyrVal 135
Db 116 LeuProPheHisLeuLeuProSerLeuLeuGluLysAsnCysLysIleIleTyrVal 135
Qy 414 GCTGCAACCCAGGATCTGGTGTCTTATATCATGTTCCACCGCTCTCTCGGACC 473
Db 414 GCTGCAACCCAGGATCTGGTGTCTTATATCATGTTCCACCGCTCTCTCGGACC 473
Qy 136 AlaArgAsnProLysAspAsnMetValSerTyrHisPheGlnArgMetAsnLysAla 155
Db 136 AlaArgAsnProLysAspAsnMetValSerTyrHisPheGlnArgMetAsnLysAla 155
Qy 474 ATGAGCTACCGAGCAGCCTTTCAGAAATCTCGCGGAGTGTATGAATGATAAGCTGGC 533
Db 474 ATGAGCTACCGAGCAGCCTTTCAGAAATCTCGCGGAGTGTATGAATGATAAGCTGGC 533
Qy 156 LeuProAlaProGlyThrTrpGluGluTyrPheGluThrPheLeuAlaGlyLysValCys 175
Db 156 LeuProAlaProGlyThrTrpGluGluTyrPheGluThrPheLeuAlaGlyLysValCys 175
Qy 534 TACGGCTCTCGTGTGAGCAGCTGAGAGTCTGGGAGCAGCAGCAGCTCGAAGCTG 593
Db 534 TACGGCTCTCGTGTGAGCAGCTGAGAGTCTGGGAGCAGCAGCAGCTCGAAGCTG 593
Qy 176 TrpGlySerTrpHisGluHisValLysGlyTrpTrpGluAlaLysAspLysHisArgIle 195
Db 176 TrpGlySerTrpHisGluHisValLysGlyTrpTrpGluAlaLysAspLysHisArgIle 195
Qy 594 CTTTCTTCAAGTATGAAGACATCGGACCTGTGTGACGATGGTGGAGCAGCTGGCC 653
Db 594 CTTTCTTCAAGTATGAAGACATCGGACCTGTGTGACGATGGTGGAGCAGCTGGCC 653
Qy 196 LeuTyrLeuPheTyrGluAspMetLysLysAsnProLysHisGluIleGlnLysLeuAla 215
Db 196 LeuTyrLeuPheTyrGluAspMetLysLysAsnProLysHisGluIleGlnLysLeuAla 215
Qy 654 AGATTCTGGGGTGTCTGTGACAGCCAGCTGGAGCCCTGAGGAGCAGCTGGCC 713
Db 654 AGATTCTGGGGTGTCTGTGACAGCCAGCTGGAGCCCTGAGGAGCAGCTGGCC 713
Qy 216 GluPheIleGlyLysLysLeuAspLysValLeuAspLysIleValHisTyrThrSer 235
Db 216 GluPheIleGlyLysLysLeuAspLysValLeuAspLysIleValHisTyrThrSer 235
Qy 714 CAGCTGTGTG-----GACCAGTGTCTGCAACCTGAGGCGCTGCCC----- 752
Db 714 CAGCTGTGTG-----GACCAGTGTCTGCAACCTGAGGCGCTGCCC----- 752
Qy 236 PheAspValMetLysGlnAsnProMetAlaAsnTyrSerSerIleProAlaGluIleMet 255
Db 236 PheAspValMetLysGlnAsnProMetAlaAsnTyrSerSerIleProAlaGluIleMet 255
Qy 753 -----GTGGCGCGGGAAGAGTTGGCTGTGGAGGACATCTTC 791
Db 753 -----GTGGCGCGGGAAGAGTTGGCTGTGGAGGACATCTTC 791
Qy 256 AspHisSerIleSerProPheMetArgLysGlyAlaValGlyAspTrpLysLysHisPhe 275
Db 256 AspHisSerIleSerProPheMetArgLysGlyAlaValGlyAspTrpLysLysHisPhe 275
Qy 792 ACCGCTCCATGATGAGAGTTTGCATGCTGTATATAACAGAGAGTGGGAAGTGTGAC 851
Db 792 ACCGCTCCATGATGAGAGTTTGCATGCTGTATATAACAGAGAGTGGGAAGTGTGAC 851
Qy 276 ThrValAlaGlnAsnGluArgPheAspGluAspTyrLysLysLysMetThrAspThrArg 295
Db 276 ThrValAlaGlnAsnGluArgPheAspGluAspTyrLysLysLysMetThrAspThrArg 295
Qy 852 CTCACGTTTGACTTT 866
Db 852 CTCACGTTTGACTTT 866
Qy 296 LeuThrPheHisPhe 300
Db 296 LeuThrPheHisPhe 300
RESULT 46
ADZ70333
ID ADZ70333 standard; protein; 302 AA.
XX
XX
XX AC ADZ70333;
XX
XX 30-JUN-2005 (first entry)
XX
XX Human protein from lung cancer marker gene SUL1C2.
XX
XX Tumor marker; lung tumor; cytostatic; neoplasm; expression;
XX DNA microarray.
XX
XX Homo sapiens.
XX
XX WO2005032495-A2.
XX
XX 14-APR-2005.
XX
XX
XX 01-OCT-2004; 2004WO-US034163.
XX
XX PF
XX PR 03-OCT-2003; 2003US-0508355P.
XX
XX (FARB ) BAYER PHARM CORP.
XX

```

```

XX
PI Taylor I, Pauloski NR, Bigwood D;
XX WPI: 2005-285325/29.
DR N-PSDB, ADZ70332.
XX
XX Providing a patient diagnosis for lung cancer comprises comparing the
FT level of expression of genes or gene products in a biological sample from
FT the patient with that from a normal individual.
XX
XX Claim 3; SEQ ID NO 18; 60pp; English.
XX
XX The invention relates to providing a patient diagnosis for lung cancer
CC comprising comparing the level of expression of genes or gene products in
CC a biological sample from the patient with the level of expression of
CC genes or gene products in a biological sample from a normal individual.
CC Also included are distinguishing between normal and disease tissues,
CC monitoring the response of a patient being treated for lung cancer by
CC administering an anti-cancer agent, identifying a compound useful for the
CC treatment of lung cancer and an array for distinguishing between normal
CC and disease tissues (comprising 2 or more probes corresponding to 2 or
CC more genes selected from any of the 200 nucleotide sequences given in the
CC specification, or 2 or more polypeptides comprising any of the 200 amino
CC acid sequences given in the specification). In providing a patient
CC diagnosis for lung cancer, one or more genes are selected from any of the
CC 200 nucleotide sequences as mentioned in the specification, or one or
CC more gene products are polypeptides selected from any of the 20 amino
CC acid sequences mentioned in the specification. The methods are useful for
CC detecting and treating lung cancer. These may also be used for designing,
CC identifying and optimizing therapeutics for cancer. The present sequence
CC represents a protein from one of the 200 lung cancer marker genes. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 302 AA;

```

```

Alignment Scores:
Pred. No.: 3,83e-33 Length: 302
Score: 449.00 Matches: 96
Percent Similarity: 55.8% Conservative: 52
Best Local Similarity: 36.2% Mismatches: 99
Query Match: 10.2% Indels: 18
DB: Gaps: 3
US-10-768-158-1 (1-2419) x ADZ70333 (1-302)

```

```

Qy 126 GAGGAGATCGCAACTTCCCGGTGCGGCGCAGCGAGCTGTGGATCGTCACTACCCCAAG 185
Db 126 GAGGAGATCGCAACTTCCCGGTGCGGCGCAGCGAGCTGTGGATCGTCACTACCCCAAG 185
Qy 36 AspLysIleTrpAsnPheGlnAlaLysProAspLeuIleSerThrTyrProLys 55
Db 36 AspLysIleTrpAsnPheGlnAlaLysProAspLeuIleSerThrTyrProLys 55
Qy 186 TCGGCGCAGCAGCTTGTCTGTCAGGAGGTGCTACTTGTGTGAGCCAGCGGCGTCCCGCAT 245
Db 186 TCGGCGCAGCAGCTTGTCTGTCAGGAGGTGCTACTTGTGTGAGCCAGCGGCGTCCCGCAT 245
Qy 56 AlaGlyThrThrTrpThrGlnGluIleValGluLeuIleGlnAsnGluGlyAspValGlu 75
Db 56 AlaGlyThrThrTrpThrGlnGluIleValGluLeuIleGlnAsnGluGlyAspValGlu 75
Qy 246 GAGATCGGCTTGATGAACATCGACGACAGCTCCCGTCTCTGGAGTACCCACAGCCG--- 302
Db 246 GAGATCGGCTTGATGAACATCGACGACAGCTCCCGTCTCTGGAGTACCCACAGCCG--- 302
Qy 76 LysSerLysArgAlaProThrHisGlnArgPheProPheLeuGluMetLysIleProSer 95
Db 76 LysSerLysArgAlaProThrHisGlnArgPheProPheLeuGluMetLysIleProSer 95
Qy 303 -----GGCTGGACATCATCAGAACTGACCTCCCGCTCATCAAGAGCCAC 353
Db 303 -----GGCTGGACATCATCAGAACTGACCTCCCGCTCATCAAGAGCCAC 353
Qy 96 LeuGlySerGlyLeuGluGlnAlaHisAlaMetProSerProArgIleLeuLysThrHis 115
Db 96 LeuGlySerGlyLeuGluGlnAlaHisAlaMetProSerProArgIleLeuLysThrHis 115
Qy 354 CTGCCCTACCGCTTCTCCCTCTGACCTCCACAAATGAGAGCTCAAGAGTTCATATATG 413
Db 354 CTGCCCTACCGCTTCTCCCTCTGACCTCCACAAATGAGAGCTCAAGAGTTCATATATG 413
Qy 116 LeuProPheHisLeuLeuProSerLeuLeuGluLysAsnCysLysIleIleTyrVal 135
Db 116 LeuProPheHisLeuLeuProSerLeuLeuGluLysAsnCysLysIleIleTyrVal 135
Qy 414 GCTGCAACCCAGGATCTGGTGTCTTATATCATGTTCCACCGCTCTCTCGGACC 473
Db 414 GCTGCAACCCAGGATCTGGTGTCTTATATCATGTTCCACCGCTCTCTCGGACC 473
Qy 136 AlaArgAsnProLysAspAsnMetValSerTyrHisPheGlnArgMetAsnLysAla 155
Db 136 AlaArgAsnProLysAspAsnMetValSerTyrHisPheGlnArgMetAsnLysAla 155
Qy 474 ATGAGCTACCGAGCAGCCTTTCAGAAATCTCGCGGAGTGTATGAATGATAAGCTGGC 533
Db 474 ATGAGCTACCGAGCAGCCTTTCAGAAATCTCGCGGAGTGTATGAATGATAAGCTGGC 533

```


Query Match:	10.2%	Indels:	58
DB:	8	Gaps:	5
US-10-768-158-1 (1-2419) x ABM84002 (1-327)			
QY	90	GGCGTGGCGTGGCGCGCTTCTGCGGGGAAGATGGAGGAGATCGCAACTTCCCGGTG	149
DB	17	GlyValProLeuIleLysTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla	36
QY	150	CGGCCACGACGATGTGGATCGTACCTTACCCCAAGTCCGACACAGCTTCTCCAGGAG	209
DB	37	ArgProAspLeuLeuIleAsnThrTyrProLysSerGlyThrThrTrpValSerGln	56
QY	210	GTGCTCTACTTGGTGGACCGGGCGTGCACCCCATGAGATCGCGCTTGATGAACATCGAC	269
DB	57	IleLeuAspMetIleTyrGlnGlyAspLeuGluLysCysAsnArgAlaProIleTyr	76
QY	270	GAGCAGTCCCGGTCTGGAGTACCCACAGCGG-----GGCCTGGACATCATC	317
DB	77	ValArgValProLeuGluValAsnAspProGlyGluProSerGlyLeuGluThrLeu	96
QY	318	AAGGAATGACTCTCTCCCGCTCATCAAGAGCCACCTGCGCTTCTCCCTCTCT	377
DB	97	LysAspThrProProArgLeuIleLysSerHisLeuProLeuAlaLeuProGln	116
QY	378	GACCTCCCAATGAGACTCCAGGTCTATATATGCTCGCAACCCCAAGGATCTGGTG	437
DB	117	ThrLeuLeuAspGlnLysValValValTyrValAlaArgAsnProLysAspValAla	136
QY	438	GTGCTTATTATGTTCCACCGTCTCTCGGACCATGAGCTACCGAGGCACCTTCAA	497
DB	137	ValSerTyrTyrHisPheHisArgMetGluLysAlaHisProGluProGlyThrTrpAsp	156
QY	498	GAATTCTCCGAGGTTTATGAATATAAGTGGGC-----	533
DB	157	SerPheLeuGluLysPheMetAlaGlyGluGlyLeuAspTyrArgLysGluGlyVal	176
QY	533	-----	533
DB	177	LysProArgGlyGlyTyrAsnValGlnGlnProCysValGlyAlaProCysProLeu	196
QY	534	-----TACGGCTCTGTTTGAGCACGTGAGGAGTCTGGGACGACCGCATG	581
DB	197	Leu***ValSerTyrGlySerTyrGlnHisValGlnGluTyrTrpGluLeuSerArg	216
QY	582	GACTCGAACGCTTTTCTCAAGTATGAACATGTCATCGGACCTGGTGGTACGATGGT	641
DB	217	ThrHisProValLeuTyrLeuPheTyrGluAspMetLysGluAsnProLysArgGluLeu	236
QY	642	GAGCAGTGGCCATCTCTGGGGTGTCTGTGCACAGGCCAGCTGGAAGCCTGACG	701
DB	237	GlnLysIleLeuGluPheValGlyArgSerLeuProGluGluThrMetAspPheMetVal	256
QY	702	GAGCACTGC-----	710
DB	257	GlnHisThrSerPheLysGluMetLysLysAsnProMetThrAsnTyrThrThrValPro	276
QY	711	CACCAGCTGGTGGACCATGTGTGCAACGCTGAGGCCCTTGCCC-----GTGGCGCGGGGAAGA	767
DB	277	GlnGluLeuMetAspHisSerIleSer-----ProPheMetArgLysGlyMet	292
QY	768	GTTCGGCTGTGAAGACATCTTACCGTCTCCATGATGACAGCTTTGACTTGGTGTAT	827
DB	293	AlaGlyAspTyrLysThrThrPheThrValAlaGlnAsnGluArgPheAspAlaAspTyr	312
QY	828	AAACAGAAATGGGAAAGTGTGACCTCACGTTT	860
DB	313	AlaGluLysMetAlaGlyCysSerLeuSerPhe	323
RESULT 49			
ABM84001			
ID	ABM84001	standard; protein; 327 AA.	
XX			
AC	ABM84001;		
XX			
DT	18-NOV-2004 (first entry)		
XX			
DE	Human diagnostic and therapeutic pprotein SEQ ID NO:4250.		
XX			
KW	gene therapy; human diagnostic and therapeutic polynucleotide; dithp.		
XX			
OS	Homo sapiens.		
XX			
PN	WO2004023973-A2.		
XX			
PD	25-MAR-2004.		
XX			
PF	12-SEP-2003; 2003WO-US028227.		
XX			
PR	12-SEP-2002; 2002US-0410259P.		
PR	12-SEP-2002; 2002US-0410260P.		
XX	(INCY-) INCYTE CORP.		
XX			
PI	Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;		
PI	Harthorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;		
PI	Mooney EM, Deleane AM, Panesar IS, Banville SC, Reddy TP;		
PI	Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Garstin EH;		
PI	Peralta CH, Anderson SB, Rioux P, Shen BJ, Wu MC, Stuve LL;		
PI	Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kircon ES;		
PI	Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;		
PI	Patury S, Shi X, Suarez CJ;		
XX			
XX	WPI; 2004-329368/30.		
DR	N-PSDB; ACN42653.		
XX			
PT	New diagnostic and therapeutic polynucleotides and polypeptides, useful		
PT	in diagnosing a condition, disease or disorder associated with human		
PT	molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or		
PT	in gene mapping.		
XX			
PS	Claim 27; Page; 190pp; English.		
XX			
CC	The invention relates to novel diagnostic and therapeutic polynucleotides		
CC	selected from one of the 2722 sequences defined in the specification. A		
CC	polynucleotide of the invention may have a use in gene therapy. The human		
CC	diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be		
CC	used to diagnose a particular condition, disease or disorder associated		
CC	with human molecules, e.g. cell proliferative disorders,		
CC	autoimmune/inflammatory disorder, developmental disorders, endocrine		
CC	disorder, neurological disorders, gastrointestinal disorders, or		
CC	infections caused by virus, bacteria, fungi or parasite. The dithp		
CC	molecules may also be used in genetic mapping in identifying individuals		
CC	from minute biological samples, in detecting single nucleotide		
CC	polymorphisms, as molecular weight markers, and for somatic or germline		
CC	gene therapy. The present sequence represents a dithp protein of the		
CC	invention. Note: The sequence data for this patent is not represented in		
CC	the printed specification, but was obtained in electronic format directly		
CC	from WIPO at www.wipo.int/pct/en/sequences/listing.htm		
XX			
SQ	Sequence 327 AA;		
Alignment Scores:			
Pred. No.:	6.2e-33	Length:	327
Score:	447.00	Matches:	101
Percent Similarity:	48.9%	Conservative:	51
Best Local Similarity:	32.5%	Mismatches:	101
Query Match:	10.2%	Indels:	58
DB:	8	Gaps:	5
US-10-768-158-1 (1-2419) x ABM84001 (1-327)			
QY	90	GGCGTGGCGTGGCGCGCTTCTGCGGGGAAGATGGAGGAGATCGCAACTTCCCGGTG	149
DB	17	GlyValProLeuIleLysTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla	36


```
Db 77 ValArgValProPheLeuGluValAsnAspProGlyGluPro***GlyLeuGluThrlieu 96
      :::::||||| ||||| |||||:::
Qy 318 AAGAACTGACCTCTCCCGCCTCATCAAGAGCACCTGCTACCGCTTCTGCTCCCTCT 377
      |||::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 97 LysAspThrProProArgLeuIleLysSerHisLeuProLeuAlaLeuLeuProGln 116
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 378 GACCTCCACATGAGACTCCCAAGCTCATCTATATGCTGCTCCCAACCCCAAGATCTGGTG 437
      |||::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 117 ThrLeuLeuAspGlnLysValValValValValAlaAlaArgAsnProLysAspValAla 136
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 438 GTGCTTATTATCAGTTCCACCGCTCTCTGCGGACCATGAGCTACCGAGCACCTTTCAA 497
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 137 ValSerTyrTyrHisPheHisArgMetClnLysAlaHisProGluProGlyThrTrpAsp 156
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 498 GAATTCGCCGAGGTTTATGAATGATAAGCTGGGC----- 533
      |||::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 157 SerPheLeuGluLysPheMetAlaGlyGluGlyLeuAspTyrArgLysGluGlyVal 176
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 533 ----- 533
Db 177 LysProArgGlyGlyTyrAsnValGlnProCysValGlyAlaProCysProLeu 196
      -----TACGGCTCTCTGTTGAGCACGTGACAGAGTTCTGGAGCACCGCATG 581
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 197 Leu***ValSerTyrGlySerTrpTyrGlnHisValGlnGluTrpTrpGluLeuSerArg 216
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 582 GACTCGAACGTCTTTTCTCAAGTATGAACATGCATCGGACCTGGTGACGATGGTG 641
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 217 ThrHisProValLeuTyrLeuPheTyrGluAspMetLysGluAsnProLysArgGluIle 236
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 642 GAGCAGCTGGCCAGATTCCTGGGGTGTCTGTGACAGGCCAGCTGGAAGCCTGACG 701
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 237 GlnLysIleLeuGluPheValGlyArgSerLeuProGluGluThrMetAspPheMetVal 256
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 702 GAGCAGTCG----- 710
Db 257 GlnHisThrSerPheLysGluMetLysLysAsnProMetThrAsnTyrThrThrValPro 276
      ::|||
Qy 711 CACGAGCTGTGGACCAAGTGTGCAACGCTGAGGCCCTGCCC---GTGGGCCGGGAAGA 767
      ::|||::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 277 GlnGluLeuMetAspHisSerIleSer-----ProPheMetArgLysGlyMet 292
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 768 GTTGGCTGTGGAAGACATCTTCCCGCTCTCCATGATGAGAAGTTTGACTTGGTGTAT 827
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 293 AlaGlyAspTrpLysThrThrPheThrValAlaGlnAsnGluArgPheAspAlaAspTyr 312
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 828 AAACAGAAGATGGGAAAGTGTGACCTCACGTTT 860
      ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 313 AlaGluLysMetAlaGlyCysSerLeuSerPhe 323
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 51
ABM84003
ID ABM84003 standard; protein; 327 AA.
XX
XX AC ABM84003;
XX
XX DT 18-NOV-2004 (first entry)
XX
XX DE Human diagnostic and therapeutic pprotein SEQ ID NO:4252.
XX
XX KW gene therapy; human diagnostic and therapeutic polynucleotide; dithp.
XX
XX OS Homo sapiens.
XX
XX FN WO2004023973-A2.
XX
XX PD 25-MAR-2004.
XX
XX PF 12-SEP-2003; 2003WO-US028227.
XX
XX PR 12-SEP-2002; 2002US-0410259P.
XX
XX PR 12-SEP-2002; 2002US-0410260P.
```

```
XX PA (INCY-) INCYTE CORP.
XX
XX Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
XX Harthorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
XX Mooney EM, Deleage AM, Panesar IS, Banville SC, Reddy TP;
XX Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
XX Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
XX Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirtan ES;
XX Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
XX Patuary S, Shi X, Suarez CJ;
XX
XX WPI: 2004-329368/30.
XX N-PSDB; ACN42655.
XX
XX New diagnostic and therapeutic polynucleotides and polypeptides, useful
XX in diagnosing a condition, disease or disorder associated with human
XX molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
XX in gene mapping.
XX
XX Claim 27; Page; 190pp; English.
XX
XX The invention relates to novel diagnostic and therapeutic polynucleotides
XX selected from one of the 2722 sequences defined in the specification. A
XX polynucleotide of the invention may have a use in gene therapy. The human
XX diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
XX used to diagnose a particular condition, disease or disorder associated
XX with human molecules, e.g. cell proliferative disorders,
XX autoimmune/inflammatory disorders, gastrointestinal disorders, or
XX infections caused by virus, bacteria, fungi or parasite. The dithp
XX molecules may also be used in genetic mapping, in identifying individuals
XX from minute biological samples, in detecting single nucleotide
XX polymorphisms, as molecular weight markers, and for somatic or germline
XX gene therapy. The present sequence represents a dithp protein of the
XX invention. Note: The sequence data for this patent is not represented in
XX the printed specification, but was obtained in electronic format directly
XX from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX
XX SQ Sequence 327 AA;
XX
XX Alignment Scores:
XX Pred. No.: 6.2e-33 Length: 327
XX Score: 447.00 Matches: 101
XX Percent Similarity: 48.9% Conservative: 51
XX Best Local Similarity: 32.5% Mismatches: 101
XX Query Match: 10.2% Indels: 58
XX DB: 8 Gaps: 5
XX
XX US-10-768-158-1 (1-2419) x ABM84003 (1-327)
Qy 90 GGCGTGGGTGCGCGCCCTTCTGCGGGGAAGATGGAGAGATCGCAACTTCCGGTG 149
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 17 GlyValProLeuIleLysTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla 36
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 150 CGGCCAGCAGCGTGTGGATCGTCACCTACCCCAAGTCCCGCACCGCTTCTCAGGAG 209
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 37 ArgProAspAspLeuLeuIleAsnThrTyrProLysSerGlyThrTrpValSerGln 56
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 210 GTGCTACTTGTGTGAGCCAGGCGCTGACCCCATGAGATCGGCTTGTGAACATCGAC 269
      ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 57 IleLeuAspMetIleTyrGlnGlyAspLeuGluLysCysAsnArgAlaProIleTyr 76
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 270 GAGCAGCTCCCGCTCTCCCGCCTCATCAAGAGCACCTGCGCCTACCGCTTCTGCCCTCT 317
      ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 77 ValArgValProPheLeuGluValAsnAspProGlyGluProSerGlyLeuGluThrlieu 96
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 318 AAGAACTGACCTCTCCCGCCTCATCAAGAGCACCTGCGCCTACCGCTTCTGCCCTCT 377
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 97 LysAspThrProProArgLeuIleLysSerHisLeuProLeuAlaLeuLeuProGln 116
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 378 GACCTCCACATGAGACTCCCAAGCTCATCTATATGCTGCTCCCAACCCCAAGATCTGGTG 437
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Db 117 ThrLeuLeuAspGlnLysValValTyrValAlaAaArgAsnProLysAspValAla 136
Qy 438 GTGCTTATTATCAGTTCACCGCTCTCTGCGGACCATGAGCTACCGAGGCACCTTTCAA 497
Db 137 ValSerTyrTyrHisPheHisArgMetGluLysAlaHisProGluProGlyThrTrpAsp 156
Qy 498 GAATTCCTCCGGAGGTTTATGAATGATAAAGCTGGGC----- 533
Db 157 SerPheLeuGluLysPheMetAlaGlyGlyLeuAspTrpArgLysGluGlyVal 176
Qy 533 ----- 533
Db 177 LysProArgGlyGlyTyrAsnValGlnProCysValGlyAlaProCysProLeu 196
Qy 534 -----TACGGCTCTGTTTGGACACGTGAGGAGTCTGGGAGCAGCGCATG 581
Db 197 Leu**ValSerTyrGlySerTrpTyrGlnHisValGlnGluTyrTrpGluLeuSerArg 216
Qy 582 GACTCGAAGCTGTTTCTCAAGTATGAACATGATCGGACCTGGTGACGATGGTG 641
Db 217 ThrHisProValLeuTyrLeuPheTyrGluAspMetLysGluAsnProLysArgGluIle 236
Qy 642 GAGCAGCTGGCCAGATTCTCGGGGTGCTGTGACAGGCCAGCTGGAAGCCCTGACG 701
Db 237 GlnLysIleLeuGluPheValGlyArgSerLeuProGluGluThrMetAspPheMetVal 256
Qy 702 GAGCACTGC----- 710
Db 257 GlnHisThrSerPheLysGluMetLysLysAsnProMetThrAsnTyrThrValPro 276
Qy 711 CACGAGCTGTGGACCACTGTGCAACGCTGAGGCCCTGCCC---GTGGGCGCGGGAAGA 767
Db 277 GlnGluLeuMetAspHisSerIleSer-----ProPheMetArgLysGlyMet 292
Qy 768 GTTGGGCTGTGGAGACATCTTCACCGTCTCCATGATGAGAGTTCCTGCTGCTAT 827
Db 293 AlaGlyAspTrpLysThrThrPheThrValAlaGlnAsnGluArgPheAspAlaAspTyr 312
Qy 828 AACACAGAAGTGGGAAAGTGTGACCTCACGTTT 860
Db 313 AlaGluLysMetAlaGlyCysSerLeuSerPhe 323

RESULT 52
ABM84004
ID ABM84004 standard; protein; 327 AA.
AC ABM84004;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human diagnostic and therapeutic pprotein SEQ ID NO:4253.
XX
KW gene therapy; human diagnostic and therapeutic polynucleotide; dithp.
XX
OS Homo sapiens.
XX
PN WO2004023973-A2.
XX
PD 25-MAR-2004.
XX
PF 12-SEP-2003; 2003MO-US028227.
XX
PR 12-SEP-2002; 2002US-0410259P.
XX
PR 12-SEP-2002; 2002US-0410260P.
XX
PA (INCY-) INCYTE CORP.
XX
PI Schmidt JP, Wright RJ, Bruns CM, Marianovic MM, Shen F;
PI Harthshorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstein EH;
PI Peralca CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LU;
PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirtton ES;
```

```
PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
PI Patury S, Shi X, Suarez CJ;
XX WPI: 2004-329368/30.
DR N-PSDB, ACN42656.
XX
PT New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT in diagnosing a condition, disease or disorder associated with human
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT in gene mapping.
XX
PS Claim 27; Page: 190pp; English.
XX
CC The invention relates to novel diagnostic and therapeutic polynucleotides
CC selected from one of the 2722 sequences defined in the specification. A
CC polynucleotide of the invention may have a use in gene therapy. The human
CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
CC used to diagnose a particular condition, disease or disorder associated
CC with human molecules, e.g. cell proliferative disorders,
CC autoimmune/inflammatory disorder, developmental disorders, endocrine
CC disorder, neurological disorders, gastrointestinal disorders, or
CC infections caused by virus, bacteria, fungi or parasite. The dithp
CC molecules may also be used in genetic mapping, in identifying individuals
CC from minute biological samples, in detecting single nucleotide
CC polymorphisms, as molecular weight markers, and for somatic or germline
CC gene therapy. The present sequence data for this patent is not represented in
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX
SQ Sequence 327 AA;
Alignment Scores:
Pred. No.: 6.2e-33 Length: 327
Score: 447.00 Matches: 101
Percent Similarity: 48.9% Conservative: 51
Best Local Similarity: 32.5% Mismatches: 101
Query Match: 10.2% Indels: 58
DB: Gaps: 5
```

```
US-10-768-158-1 (1-2419) x ABM84004 (1-327)
Qy 90 GGGCTGGCTGCGCGCTTCTGCGGGAAGATGGAGGATCGCAACTTCCCGGTG 149
Db 17 GlyValProLeuIleLysTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla 36
Qy 150 CGGCCAGCAGCAGTGTGGATCGTCACCTACCCAGTCCGCGACCCAGCTTGTCTGAGGAG 209
Db 37 ArgProAspAspLeuLeuIleAsnThrTyrProLysSerGlyThrThrTrpValSerGln 56
Qy 210 GTGCTCTACTTGTGAGCAGGCGCTGACCCGATGAGATCGGCTTGTGAACTCGAC 269
Db 57 IleLeuAspMetIleTyrGlnGlyAspLeuGluLysCysAsnArgAlaProIleTyr 76
Qy 270 GAGCAGCTCCCGTCTCTGGAGTACCCACAGCCG-----GGCTGGACATCATC 317
Db 77 ValArgValProPheLeuGluValAsnAspProGlyGluProSerGlyLeuGluThrLeu 96
Qy 318 AAGGAAGTACCTCTCCCGCTCATCAAGAGCCACCTGCCCTTCTGCTGCTCTCT 377
Db 97 LysAspThrProProArgLeuIleLysSerHisLeuProLeuAlaLeuLeuProGln 116
Qy 378 GACCTCCACAATGAGACTCCAAAGGTCTATATGCTCGCAACCCCAAGAGTCTGGT 437
Db 117 ThrLeuLeuAspGlnLysValValTyrValAlaArgAsnProLysAspValAla 136
Qy 438 GTGCTTATTATCAGTTCACCGCTCTCTCGGAGCATGAGTACCGAGGCACCTTTCAA 497
Db 137 ValSerTyrTyrHisPheHisArgMetGluLysAlaHisProGluProGlyThrTrpAsp 156
Qy 498 GAATTCCTCCGGAGGTTTATGAATGATGATAAGCTGGGC----- 533
Db 157 SerPheLeuGluLysPheMetAlaGlyGlyLeuAspTrpArgLysGluGlyVal 176
```

QY 533 ----- 533

Db 177 LysProArgGlyGlyTyrAsnValGlnGlnProCysValGlyAlaProCysProLeu 196

QY 534 -----TACGGCTCTGTTGAGCACGTGCAGAGAGTTCCTGGAGCACCGCATG 581

Db 197 Leu***ValSerTyrGlySerTrpTyrGlnHisValGlnGlnTyrTrpGluLeuSerArg 216

QY 582 GACTCGAACGTCTTTTCTCAAGATGACAGACATGATCGGACCTGGTGACATGGTG 641

Db 217 ThrHisProValLeuTyrLeuPheTyrGluAspMetLysGluAsnProLysArgGluIle 236

QY 642 GAGCAGCTGGCCATCTCTGGGGGTCTCTGTGCACAGGCCAGCTGGAAGCCCTGACG 701

Db 237 GlnLysIleLeuGluPheValGlyArgSerLeuProGluGluThrMetAspPheMetVal 256

QY 702 GAGCACTGC----- 710

Db 257 GlnHisThrSerPheLysGluMetLysLysAsnProMetThrAsnTyrThrValPro 276

QY 711 CACGAGCTGGGACCGAGTGTGCAACGCTGAGGCCCTGCCC---GTGGGCCGGGGAAGA 767

Db 277 GlnGluLeuMetAspHisSerIleSer-----ProPheMetArgLysGlyMet 292

QY 768 GTTGGGCTGTGGAAGGACATCTTCACCGTCTCCATGAATGAGAAGTTTGACTTGGTGAT 827

Db 293 AlaGlyAspTrpLysThrThrPheThrValAlaGlnAsnGluArgPheAspAlaAspTyr 312

QY 828 AAACAGAAGTGGGAAGTGTGACCTCACGTTT 860

Db 313 AlaGluLysMetAlaGlyCysSerLeuSerPhe 323

RESULT 53

AD117319

ID AD117319 standard; protein; 269 AA.

XX AC AD117319;

XX DT 15-APR-2004 (first entry)

XX DE Polypeptide homologous to a human NOVX domain SeqID 855.

XX KW NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes; inflammation;

XX KW autoimmune disorder; allergy; blood disorder;

XX KW acquired immunodeficiency syndrome; AIDS; obesity; asthma;

XX KW immunoglobulin (Ig)A nephropathy; cirrhosis; arthritis;

XX KW Alzheimer's disease; infection; str.

XX OS Unidentified.

XX PN WO200268649-A2.

XX PD 06-SEP-2002.

XX PF 31-JAN-2002; 2002WO-US002785.

XX PR 31-JAN-2001; 2001US-02653395P.

XX PR 31-JAN-2001; 2001US-0265412P.

XX PR 31-JAN-2001; 2001US-0265514P.

XX PR 31-JAN-2001; 2001US-0265517P.

XX PR 02-FEB-2001; 2001US-0266406P.

XX PR 05-FEB-2001; 2001US-0266767P.

XX PR 07-FEB-2001; 2001US-0266975P.

XX PR 07-FEB-2001; 2001US-0267057P.

XX PR 08-FEB-2001; 2001US-0267459P.

XX PR 09-FEB-2001; 2001US-0267823P.

XX PR 15-FEB-2001; 2001US-0268974P.

XX PR 26-FEB-2001; 2001US-0271664P.

XX PR 27-FEB-2001; 2001US-0271839P.

XX PR 27-FEB-2001; 2001US-0271855P.

XX PR 02-MAR-2001; 2001US-0272788P.

XX PR 02-MAR-2001; 2001US-0273046P.

PR 14-MAR-2001; 2001US-0275925P.

PR 14-MAR-2001; 2001US-0275947P.

PR 14-MAR-2001; 2001US-0275950P.

PR 15-MAR-2001; 2001US-0275989P.

PR 15-MAR-2001; 2001US-0276448P.

PR 15-MAR-2001; 2001US-0276450P.

PR 16-MAR-2001; 2001US-0276397P.

PR 16-MAR-2001; 2001US-0276768P.

PR 20-MAR-2001; 2001US-0278652P.

PR 26-MAR-2001; 2001US-0278775P.

PR 26-MAR-2001; 2001US-0278778P.

PR 29-MAR-2001; 2001US-0279882P.

PR 29-MAR-2001; 2001US-0279884P.

PR 30-MAR-2001; 2001US-0280147P.

PR 11-APR-2001; 2001US-0282992P.

PR 11-APR-2001; 2001US-0283083P.

PR 20-APR-2001; 2001US-0285133P.

PR 23-APR-2001; 2001US-0285749P.

PR 03-MAY-2001; 2001US-0288327P.

PR 03-MAY-2001; 2001US-0288504P.

PR 29-MAY-2001; 2001US-0294047P.

PR 30-MAY-2001; 2001US-0294473P.

PR 08-JUN-2001; 2001US-0296964P.

PR 18-JUN-2001; 2001US-0298959P.

PR 19-JUN-2001; 2001US-0299324P.

PR 13-AUG-2001; 2001US-0312020P.

PR 16-AUG-2001; 2001US-0312889P.

PR 16-AUG-2001; 2001US-0312908P.

PR 21-AUG-2001; 2001US-0313390P.

PR 28-AUG-2001; 2001US-0315470P.

PR 31-AUG-2001; 2001US-0316447P.

PR 07-SEP-2001; 2001US-0318115P.

PR 07-SEP-2001; 2001US-0318118P.

PR 12-SEP-2001; 2001US-0318740P.

PR 19-SEP-2001; 2001US-0323379P.

PR 18-OCT-2001; 2001US-0330245P.

PR 18-OCT-2001; 2001US-0330308P.

PR 14-NOV-2001; 2001US-0332701P.

XX XX

PA (CURA-) CURAGEN CORP.

XX PI Tchernev VT, Spytek KA, Zehrusen BD, Patturajan M, Shinkets RA;

XX PI Li L, Gangolli EA, Padigaru M, Anderson DW, Rastelli L, Miller CE;

PI Gerlach VL, Taupier RJ, Gusev VY, Colman SD, Wolenc AR, Pena CE;

PI Furtak K, Grosse WM, Alsobrook JP, Lepley DM, Rieger DK, Burgess CE;

XX XX

DR WPI; 2002-706998/76.

XX XX

PT New NOVX polypeptides and nucleic acids, useful for preventing or

PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,

PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or

PT pharmacogenomics.

XX XX

PS Disclosure; SEQ ID NO 855; 1498pp; English.

XX XX

CC This invention relates to a novel nucleic acids, and encoded polypeptides

CC thereof, which have properties related to the stimulation of biochemical

CC or physiological responses in a cell, tissue, organ or organism.

CC Specifically, it refers to the use of biologically active fragments for

CC diagnostic and prognostic assays and furthermore in the treatment of

CC diverse pathological conditions. The present invention describes novel

CC human and murine NOVX proteins, as well as methods to modulate their

CC expression using antisense oligos, ribozymes and peptide nucleic acids.

CC The NOVX polypeptides, polynucleotides and antibodies are useful in

CC treating or preventing NOVX-associated disorders, e.g. cardiomyopathy,

CC atherosclerosis, cancer and diabetes. Furthermore, they may be used in

CC treating or preventing diseases such as inflammation, autoimmune

CC disorders, allergies, blood disorders, acquired immunodeficiency syndrome

CC (AIDS), obesity, asthma, immunoglobulin (Ig)A nephropathy, cirrhosis,

CC arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy

CC and epilepsy. Accordingly, these molecules have many activities including

CC cytostatic, cardiant, antiinflammatory, immunosuppressive, antiallergic,

CC haemostatic, anti-HIV, antidiabetic, antiarteriosclerotic, anorectic,

CC


```
Db      6 AspTyrTyrGluLysPheGluGluValHisGlyIleLeuMetTyrLysAspPheVallys 25
Qy      120 AAGATGGAGGAGATCGCAACTTCCCGGTGGCGCCACGACGCTGTGGATCGTCACCTAC 179
Db      26 TyrTrpAspAsnValGluAlaPheGlnAlaArgProAspAspLeuValIleAlaThrTyr 45
Qy      180 CCCAAGTCGGCAGCAGCAGCTTCTGCGAGAGTGTCTACTTGGTGAGCCAGGGCGCTGAC 239
Db      46 ProlysSerGlyThrTrpValSerGluIleValTyrMetIleTyrLysGluGlyAsp 65
Qy      240 CCCGATGAGTCGGCTTGATGAACATCGACGAGCAGCTCCCGTCTGGAGTACCCACAG 299
Db      66 ValGluLysCysLysGluAspValIlePheAsnArgIleProPheLeuGluCysArglys 85
Qy      300 CCG-----GGCCTGGACATCATCAGGAACCTGACCTCCCGGCTCATCAAG 347
Db      86 GluAsnLeuMetAsnGlyValLysGlnLeuAspGluMetAsnSerProArgIleVallys 105
Qy      348 AGCCACCTGCGCTACCGCTTCTGCGCTCTGACCTCCCAATGGAGACTCCAAGTCTATC 407
Db      106 ThrHisLeuProGluLeuLeuProAlaSerPheTrpGluLysAspCysLysIleIle 125
Qy      408 TATATGGCTCGCAACCCCAAGGATCTGCTGTGTCTTATTATTCAGTTCCACCGCTCTCTG 467
Db      126 TyrLeuCysArgAsnAlaLysAspValAlaValSerPheTyrTyrPheLeuMetVal 145
Qy      468 CGGACCATGAGCTACCGAGGCACCTTTCAAGATTCTGCGGAGTTCGCGAGGTTTATGATGAAG 527
Db      146 AlaGlyHisProAsnProGlySerPheProGluPheValGluLysPheMetGlnGlyGln 165
Qy      528 CTGGGCTCGGCTCTGCTGTTTGGACAGCTGACGAGTTCCTGGGAGCACCCGATCGACTCG 587
Db      166 ValProTyrGlySerTrpTyrLysHisValLysSerTrpTrpGluLysGlyLysSerPro 185
Qy      588 AACGTGCTTTTCTCAAGTATGAAGACATGCATCGGACCTGTCGACGATGGTGGAGCAG 647

Db      186 ArgValLeuPheLeuPheTyrGluAspLeuLysGluAspIleArgLysGluValIleLys 205
Qy      648 CTGCCAGATCTCTGGGGGTCTCTGTGACAAAGCCAGCTGGAAGCCCTGACGGAGCAC 707
Db      206 LeuIleHisPheLeuGluAtrGlyProSerGluGluLeuValAspArgIleIleHisHis 225
Qy      708 TGCAC-----CAGCTGCTGGACAC 728
Db      226 ThrSerPheGlnGluMetLysAsnAsnProSerThrAsnTyrThrThrLeuProAspGlu 245
Qy      729 TGCTGCAACGCTGAGGCCCTGCC--GTGGCGCGGGAAGAGTTGGGCTGTGGAAGGAC 785
Db      246 IleMetAsnGlnLysLeuSerProPheMetArgLysGlyIleThrGlyAspTrpLysAsn 265
Qy      786 ATCTTCCCGTCTCCATGAATGAGAAGTTTGACTTGGTGTATAAACAGAAGATGGGAAAG 845
Db      266 HisPheThrValAlaLeuAsnGluLysPheAspLysHisTyrGluGlnMetLysGlu 285
Qy      846 TGTGACCTCAGTTT 860
Db      286 SerThrLeuLysPhe 290

RESULT 58
AAW44247
ID    AAW44247 standard; protein; 294 AA.
XX
AC    AAW44247;
XX
DT    15-MAY-1998 (first entry)
XX
DE    Human oestrogen sulphotransferase.
XX
KW    Human; oestrogen sulphotransferase; estrogen sulphotransferase; liver EST;
XX    biotransformation.
OS    Homo sapiens.
XX
```

```
PN      US5714594-A.
XX
PD      03-FEB-1998.
XX
PF      18-OCT-1994; 94US-00325562.
XX
PR      18-OCT-1994; 94US-00325562.
XX
PA      (MAYO-) MAYO FOUNDATION.
XX
PI      Wood TC, Weinshilboum RM, Aksoy IA;
XX
DR      WPI; 1998-144284/13.
XX
PT      N-PSDB; AAV12466.
XX
SQ      DNA encoding human oestrogen sulphotransferase - useful for research into
        oestrogen biotransformation.
XX
PS      Claim 1; Col 21-24; 25pp; English.
XX
CC      The present sequence represents a human oestrogen sulphotransferase (EST)
        protein. The isolation and expression of cDNA which encodes human liver
        CC      EST enables the biotransformation of oestrogens to be studied
        CC      Sequence 294 AA;
        SQ

Alignment Scores:
Pred. No.: 1.02e-31 Length: 294
Score: 434.00 Matches: 95
Percent Similarity: 53.0% Conservative: 56
Best Local Similarity: 33.3% Mismatches: 116
Query Match: 9.9% Indels: 18
DB: 2 Gaps: 3

US-10-768-158-1 (1-2419) x AAW44247 (1-294)
Qy      60 GAGTTCGAGACCAAGTACTTCGAGTTCATGGCTGCGGTGCGGCTTCTGCGCGGG 119
Db      6 AspTyrTyrGluLysPheGluGluValHisGlyIleLeuMetTyrLysAspPheVallys 25
Qy      120 AAGATGAGGAGATCGCAACTTCCCGTGGCGCCACGACGCTGATCGTCACCTAC 179
Db      26 TyrTrpAspAsnValGluAlaPheGlnAlaArgProAspAspLeuValIleAlaThrTyr 45
Qy      180 CCCAAGTCGGCAGCAGCAGCTTGTCTGACGAGTGTCTACTTGGTGAGCCAGGGCGCTGAC 239
Db      46 ProlysSerGlyThrTrpValSerGluIleValTyrMetIleTyrLysGluGlyAsp 65
Qy      240 CCCGATGAGTCGGCTTGATGAACATCGACGAGCAGCTCCCGTCTGGAGTACCCACAG 299
Db      66 ValGluLysCysLysGluAspValIlePheAsnArgIleProPheLeuGluCysArglys 85
Qy      300 CCG-----GGCCTGGACATCATCAGGAACCTGACCTCCCGGCTCATCAAG 347
Db      86 GluAsnLeuMetAsnGlyValLysGlnLeuAspGluMetAsnSerProArgIleVallys 105
Qy      348 AGCCACCTGCGCTACCGCTTCTGCGCTCTGACCTCCCAATGGAGACTCCAAGTCTATC 407
Db      106 ThrHisLeuProGluLeuLeuProAlaSerPheTrpGluLysAspCysLysIleIle 125
Qy      408 TATATGGCTCGCAACCCCAAGGATCTGCTGTGTCTTATTATTCAGTTCCACCGCTCTCTG 467
Db      126 TyrLeuCysArgAsnAlaLysAspValAlaValSerPheTyrTyrPheLeuMetVal 145
Qy      468 CGGACCATGAGCTACCGAGGCACCTTTCAAGATTCTGCGGAGTTCGCGAGGTTTATGATGAAG 527
Db      146 AlaGlyHisProAsnProGlySerPheProGluPheValGluLysPheMetGlnGlyGln 165
Qy      528 CTGGGCTCGGCTCTGCTGTTTGGACAGCTGACGAGTTCCTGGGAGCACCCGATCGACTCG 587
Db      166 ValProTyrGlySerTrpTyrLysHisValLysSerTrpTrpGluLysGlyLysSerPro 185
Qy      588 AACGTGCTTTTCTCAAGTATGAAGACATGCATCGGACCTGTCGACGATGGTGGAGCAG 647
```


Db 206 LeuIleHisPheLeuGluArgLysProSerGluGluLeuValAspArgIleIleHisHis 225
Qy 708 TGCCAC-----CAGCTGGTGACAC 728
Db 226 ThrSerPheGlnGluMetLysAsnAsnProSerThrAsnTyrThrLeuProAspGlu 245
Qy 729 TGCTGCAACGCTGAGCCCTGCC---GTGGGCGGGAGAGTTGGCTGTGAAGAC 785
Db 246 IleMetAsnGlnLysLeuSerProPheMetArgLysGlyIleThrGlyAspTrpLysAsn 265
Qy 786 ATCTTCACCGTCCATGACAGTTTGACTTGGTGATATAACACAGACATGGGAAG 845
Db 266 HisPheThrValAlaLeuAsnGluLysPheAspLysHisTyrGluGlnGlnMetLysGlu 285
Qy 846 TGTGACCTCACGTTT 860
Db 286 SerThrLeuLysPhe 290
RESULT 60
ADII17049
ID ADII17049 standard; protein; 307 AA.
XX
AC ADII17049;
XX
DT 15-APR-2004 (first entry)
XX
DE Chicken NOVX protein homologue SeqID 585.
XX
KW chicken; NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes;
KW inflammation; autoimmune disorder; allergy; blood disorder;
KW acquired immunodeficiency syndrome; AIDS; obesity; asthma;
KW immunoglobulin (Ig)A nephropathy; cirrhosis; arthritis;
KW Alzheimer's disease; infection; s.
XX
OS Gallus gallus.
XX
PN W0200268649-A2.
XX
PD 06-SEP-2002.
XX
PF 31-JAN-2002; 2002WO-US002785.
XX
PR 31-JAN-2001; 2001US-0265395P.
PR 31-JAN-2001; 2001US-0265412P.
PR 31-JAN-2001; 2001US-0265514P.
PR 31-JAN-2001; 2001US-0265517P.
PR 02-FEB-2001; 2001US-0266406P.
PR 05-FEB-2001; 2001US-0266787P.
PR 07-FEB-2001; 2001US-0266975P.
PR 08-FEB-2001; 2001US-0267057P.
PR 09-FEB-2001; 2001US-0267823P.
PR 15-FEB-2001; 2001US-0268974P.
PR 26-FEB-2001; 2001US-0271684P.
PR 27-FEB-2001; 2001US-0271839P.
PR 27-FEB-2001; 2001US-0271855P.
PR 02-MAR-2001; 2001US-0272788P.
PR 02-MAR-2001; 2001US-0273046P.
PR 14-MAR-2001; 2001US-0275925P.
PR 14-MAR-2001; 2001US-0275947P.
PR 14-MAR-2001; 2001US-0275950P.
PR 15-MAR-2001; 2001US-0275989P.
PR 15-MAR-2001; 2001US-0276448P.
PR 15-MAR-2001; 2001US-0276450P.
PR 16-MAR-2001; 2001US-0276439P.
PR 16-MAR-2001; 2001US-0276788P.
PR 20-MAR-2001; 2001US-0278652P.
PR 26-MAR-2001; 2001US-0278775P.
PR 26-MAR-2001; 2001US-0278778P.
PR 29-MAR-2001; 2001US-0279882P.
PR 29-MAR-2001; 2001US-0279884P.
PR 30-MAR-2001; 2001US-0280147P.
PR 11-APR-2001; 2001US-0282992P.

PR 11-APR-2001; 2001US-0283083P.
PR 20-APR-2001; 2001US-0285133P.
PR 23-APR-2001; 2001US-0285749P.
PR 03-MAY-2001; 2001US-0288327P.
PR 03-MAY-2001; 2001US-0288504P.
PR 29-MAY-2001; 2001US-0294047P.
PR 30-MAY-2001; 2001US-0294473P.
PR 08-JUN-2001; 2001US-0296964P.
PR 18-JUN-2001; 2001US-0298959P.
PR 19-JUN-2001; 2001US-0299324P.
PR 13-AUG-2001; 2001US-0312020P.
PR 16-AUG-2001; 2001US-0312889P.
PR 21-AUG-2001; 2001US-0313908P.
PR 28-AUG-2001; 2001US-0315470P.
PR 31-AUG-2001; 2001US-0316447P.
PR 07-SEP-2001; 2001US-0318115P.
PR 07-SEP-2001; 2001US-0318118P.
PR 12-SEP-2001; 2001US-0318740P.
PR 19-SEP-2001; 2001US-0323379P.
PR 18-OCT-2001; 2001US-0330245P.
PR 18-OCT-2001; 2001US-0330308P.
PR 14-NOV-2001; 2001US-0332701P.
XX
PA (CURA-) CURAGEN CORP.
XX
XX Tchernev VT, Spytek KA, Zerhusen BD, Patturajan M, Shinkets RA;
PI Li L, Gangolli EA, Padigaru M, Anderson DM, Rastelli L, Miller CE;
PI Garlach VL, Taupier RJ, Gusev VY, Colman SD, Wolenc AR, Pena CEA;
PI Furtak K, Grosse WM, Alsobrook JP, Lepley DM, Rieger DK, Burgess CE;
XX
DR WPI; 2002-706998/76.
XX
XX New NOVX polypeptides and nucleic acids, useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy, typing or
PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
PT pharmacogenomics.
XX
PS Disclosure; SEQ ID NO 585; 1498pp; English.
XX
XX This invention relates to a novel nucleic acids, and encoded polypeptides
CC thereof, which have properties related to the stimulation of biochemical
CC or physiological responses in a cell, tissue, organ or organism.
CC Specifically, it refers to the use of biologically active fragments for
CC diagnostic and prognostic assays and furthermore in the treatment of
CC diverse pathological conditions. The present invention describes novel
CC human and murine NOVX proteins, as well as methods to modulate their
CC expression using antisense oligos, ribozymes and peptide nucleic acids.
CC The NOVX polypeptides, polynucleotides and antibodies are useful in
CC treating or preventing NOVX-associated disorders, e.g. cardiomyopathy,
CC atherosclerosis, cancer and diabetes. Furthermore, they may be used in
CC treating or preventing diseases such as inflammation, autoimmune
CC disorders, allergies, blood disorders, acquired immunodeficiency syndrome
CC (AIDS), obesity, asthma, immunoglobulin (Ig)A nephropathy, cirrhosis,
CC arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy
CC and epilepsy. Accordingly, these molecules have many activities including
CC cytostatic, cardiac, antiinflammatory, immunosuppressive, antiallergic,
CC haemostatic, anti-HIV, antidiabetic, antiatherosclerotic, anorectic,
CC antiasthmatic, nephrotropic, antiarthritic, hepatotropic,
CC neuroprotective, nootropic, antibacterial, virucide, antiparasitic,
CC relaxant and anticonvulsant. In addition, they are useful in screening
CC assays to identify small molecules that modulate or inhibit, for example,
CC neurogenesis, wound healing and angiogenesis. The nucleic acids are also
CC used as in chromosome mapping, tissue typing, preventive medicine and
CC pharmacogenomics. This polypeptide is a homologue of a human NOVX protein
CC of the invention.
XX
SQ Sequence 307 AA;
Alignment Scores:
Pred. No.: 3.89e-31 Length: 307
Score: 428.00 Matches: 87
Percent Similarity: 55.4% Conservative: 61

Best Local Similarity: 32.6%			Mismatches: 97		
Query Match: 9.7%			Indels: 22		
			Gaps: 2		
US-10-768-158-1 (1-2419) x AD117049 (1-307)					
Qy	126	GAGGAGATCGCAACTTCCCGGTGGCGGCCAGGACGAGTGTGGATCGTCACTACCCCAAG	185		
Db	37	AspGlnValTrpAsnPheLysAlaArgProAspAspLeuValAlaThrTyrAlaLys	56		
Qy	186	TCGGCCACCACTTCTGCAGAGGTGGTCTACTTGGTGAGCCAGGGCGCTGACCCCGAT	245		
Db	57	AlaGlyThrThrTyrThrGlnGluLeuValAspMetIleGlnAsnGlyAspIleGlu	76		
Qy	246	GAGATCGGCTTGATGAACATCGACGAGCATCCCGTCTGGAGTACCCACACGCG---	302		
Db	77	LysCysArgArgAlaSerThrTyrLysArgHisProPheLeuGluTyrTyrIleProAsp	96		
Qy	303	-----GGCCTGGACATCATCAAGGAACGACCTCTCCCGCGCTC	341		
Db	97	SerSerProLeuGlyTyrSerGlyLeuLysLeuAlaGluAlaMetProSerProArgThr	116		
Qy	342	ATCAAGACCCACCTGCCCTACCGCTTCTGCCCTCTGACCTCCACAATGGAGACTCCAAG	401		
Db	117	MetLysThrHisLeuProValGlnLeuValProProSerPheTrpGluGlnAsnCysLys	136		
Qy	402	GTCATCTATATGGCTCGCAACCCCAAGGATCTGGTGTCTTATTATCAGTTTCCACCGC	461		
Db	137	IleIleTyrValAlaArgAsnAlaLysAspAsnLeuValSerTyrTyrHisPheHisArg	156		
Qy	462	TCTCTGGGACCATGAGCTACCGAGGCACCTTTCAGAAATTCCTCGCGGAGTTTATGAAT	521		
Db	157	MetAsnLysValLeuProAspProGlyThrIleGluPheThrGluLysPheMetAsn	176		
Qy	522	GATAAGCTGGGCTACGGCTCTCTGGTTTGAGCACGTGCAGGAGTTCTGGGAGCACCGCATG	581		
Db	177	GlyGluValLeuTyrGlySerTyrTrpAspHisValLysGlyTyrTrpLysAlaLysAsp	196		
Qy	582	GACTCGAACGTCTTTTCTCAAGTATGAACATGCATCGGACCTGGTGACATGGTGATGGT	641		
Db	197	LysHisArgIleLeuTyrLeuPheTyrGluAspMetLysGluAsnProLysArgGluIle	216		
Qy	642	GAGCAGCTGGCCAGATCTCTGGGGTGTCTCTGTGACAGGCCACGCTG-----	689		
Db	217	GlnLysIleMetLysPheLeuGluLysAspLeuAspGluGluValLeuAsnLysIleIle	236		
Qy	690	-----GAAAGCCCTGACGAGCAGCTGCCACGACGCTG	719		
Db	237	TyrAsnThrSerPheGluIleMetLysAspAsnProMetThrAsnTyrThrLysAspPhe	256		
Qy	720	GTGACACGAGTCTCAACGCTGAGCCCTGCGCTGGCGCGGGAAGAGTTGGCTGGTGG	779		
Db	257	ValGlyValMetAspHisSerValSerProPheMetArgLysGlySerValGlyAspTrp	276		
Qy	780	AAGGACATCTTCCAGCTCTCCATGAATGAGAAGTTTGAATGGTGTATAACACAGAGATG	839		
Db	277	LysAsnTyrPheThrValAlaLeuAsnLysLysPheAspGlnAspTyrLysLysLysMet	296		
Qy	840	GGAAAGTGTGACCTCAGCTTT	860		
Db	297	AlaAspThrSerLeuValPhe	303		
RESULT 61					
AAE12209					
ID	AAE12209 standard; protein; 304 AA.				
XX	AAE12209;				
XX	18-DEC-2001 (first entry)				
DT	Human ST drug-metabolising protein 2 encoded by DNA transcript 2.				
DE	Human; sulfotransferase; drug-metabolising enzyme mediated disorder; ST;				
XX					
XX					

KW	DME; toxicity study; gene therapy.				
XX	Homo sapiens.				
OS					
XX					
PH	Location/Qualifiers				
FT	Modified-site	173..175			
FT	Modified-site	/note= "Protein kinase C phosphorylation site"			
FT	Modified-site	180..183			
FT	Modified-site	/note= "Casein kinase II phosphorylation site"			
FT	Modified-site	236..239			
FT	Modified-site	/note= "Casein kinase II phosphorylation site"			
FT	Modified-site	248..251			
FT	Modified-site	/note= "Asn is N-glycosylated"			
FT	Modified-site	255..258			
FT	Modified-site	/note= "Casein kinase II phosphorylation site"			
FT	Modified-site	295..300			
FT	Modified-site	/note= "N-myristoylation site"			
FT	Modified-site	299..301			
FT	Modified-site	/note= "Protein kinase C phosphorylation site"			
XX	WO200172977-A2.				
PN					
XX					
XX	04-OCT-2001.				
PD					
XX	27-MAR-2001; 2001WO-US009659.				
PF					
PF	27-MAR-2000; 2000US-0192408P.				
XX	20-JUN-2000; 2000US-0212725P.				
PR	03-JUN-2000; 2000US-00609816.				
PR					
XX	(PEKE) PE CORP NY.				
PA					
XX	Woodage T, Wei M, Kodira C, Di Francesco V, Beasley EM;				
PI					
XX	WPI; 2001-616484/71.				
DR	N-PSDB; AAD19777.				
DR					
XX	New polypeptides useful in labeling reagents with tagged sulfur and identifying modulators of the protein comprises proteins related to sulfotransferase drug-metabolizing enzyme subfamily.				
PT					
PT	Claim 1; Fig 2; 83pp; English.				
XX					
PS	The invention relates to sulfotransferase (ST) drug-metabolising enzyme (DME) peptides, proteins and nucleic acid molecules encoding them.				
CC	Sequences of the invention are useful in labelling reagents with tagged sulphur, for the development of human therapeutics, as query sequences to perform a search against sequence databases, for identifying other family members or related sequences, to raise antibodies or to elicit another immune response, as a reagent in assays designed to quantitatively determine levels of the protein in biological fluids, to metabolise compounds for toxicity studies, and as markers for tissues in which the corresponding protein is preferentially expressed. Polypeptides of the invention is also useful in diagnosing a disease, predisposition to a disease mediated by the peptide, in pharmacogenomic analysis and treating a disorder characterised by absence of inappropriate or unwanted expression of the protein. Polynucleotides of the invention are useful for constructing recombinant vectors and host cells, expressing antigenic portions of the proteins, in genotyping and mapping experiments, for designing ribozymes, for monitoring the effectiveness of modulating compounds on the expression or activity of the protein, in diagnostic assays, antisense constructs to control protein expression and in gene therapy. The present sequence is ST drug-metabolising protein 2 encoded by DNA transcript 2				
CC					
XX	SQ Sequence 304 AA;				
Alignment Scores:					
Pred No.:	1.45e-30	Length:	304		
Score:	422.00	Matches:	97		
Percent Similarity:	52.0%	Conservative:	59		
Best Local Similarity:	32.3%	Mismatches:	124		
Query Match:	9.6%	Indels:	20		

```
DB: 4 Gaps: 5
US-10-768-158-1 (1-2419) x ABE12209 (1-304)
Qy 21 ATGCGGAGAGCGAGCGGCGAGACCCAGACACCCCGGGGAGTTCGAG---AGCAAGTAC 77
Db 1 MetAlaLysIleGluLeuAsnAlaProThrMetGluLysLysProGluLeuPheAsnIle 20
Qy 78 TTCAGTTCATGGGTGGCGTGGCGCCCTTTCGCGCGGGAAGATGAGGAGATCGCC 137
Db 21 MetGluValAspGlyValProThrLeuLeuSerLysGluTrpTrpGluLysValCys 40
Qy 138 AACTTCCCGTGGCGCCAGCGAGTGGGATCGTCACTACCCCAAGTCCGCGCACCGC 197
Db 41 AsnPheGlnAlaLysProAspAspLeuAlaThrTrpProLysSerGlyThrThr 60
Qy 198 TTGCTGCAGAGGTGGTCTACTTGGTGAGCCAGCGCGCTGACCCCGATGAGATCGGCTTG 257
Db 61 TrpMetHisGluIleLeuAspMetIleLeuAsnAspGlyAspValGluLysCysLysArg 80
Qy 258 ATGAACATCGACGAGACTCCCGGTCTCGAG-----TACCCA-----CAGCG 302
Db 81 AlaGlnThrLeuAspArgHisAlaPheLeuGluLeuLysPheProHisLysGluLysPro 100
Qy 303 GGCGTGGACATCATCAAGGAATGACCTCCCGCTCATCAAGAGCCACTGCCCTAC 362
Db 101 AspLeuGluPheValLeuGluMetSerSerProGlnLeuIleLysThrHisLeuProSer 120
Qy 363 CGCTTTCTGCCTCTGACCTCCACAATGGAGACTCCCAAGGTCACTATATGGCTCGCAAC 422
Db 121 HisLeuLeuProProSerIleTrpLysGluAsnCysIleValTrpValAlaArgAsn 140
Qy 423 CCCAAGATCTGGTGGTCTTATATAGTTCCACCGCTCTCTCGGACCATAGGTAC 482
Db 141 ProlysAspCysLeuValSerTy-Tyr-HisPheHisArgMetAlaSerPheMetProAsp 160
Qy 483 CGAGGCACCTTCAAGATCTCGCGGAGTTTATGATGATAGCTGGGTACGGTCC 542
Db 161 ProGlnAsnLeuGluGluPheTy-TyrGluLysPheMetSerGlyLysValValGlyLys 180
Qy 543 TGGTTTGAGCAGTGCAGGAGTCTCGGAGCACCAGCATGGACTCGAAGCTGCTTTTCTC 602
Db 181 TrpPheAspHisValLysGlyTrpTrpAlaAlaLysAspMetHisArgIleLeuTyrl 200
Qy 603 AAGTATGAAGACATGCATCGGACCTGCTGACGATGTTGGAGCAGCTGGCCAGATCTCTG 662
Db 201 PheTyrGluAspIleLysLysAspProlysArgGluIleGluLysIleLeuLysPheLeu 220
Qy 563 GGGGTGCTCTGTGACAGGCCACCTGGAAGCCCTGACGAGCAGTCCACCATCGGTG 722
Db 221 GluLysAspIleSerGluGluIleLeuAsnLysIleIleTyHisThrSerPheAspVal 240
Qy 723 -----GACCATGTGTCAGACGCTGAGGCCCTGCGCGTGGGC----- 758
Db 241 MetLysGlnAsnProMetThrAsnTy-TyrThrLeuProThrSerIleMetAspHisSer 260
Qy 759 -----CGGGGAAGAGTGTGGGTGGAGGACATCTTCACCGTCTCC 800
Db 261 IleSerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTy-TyrPheThrValAla 280
Qy 801 ATGAATGAGAAGTTGACTTGGTGTATATACAGAAAGATGGAAAGTGTGACCTCAGCTTT 860
Db 281 GlnAsnGluGluPheAspLysAspTy-TyrGlnLysLysMetAlaGlySerThrLeuThrPhe 300
RESULT 62
ABBB1792
ID ABB81792 standard; protein; 304 AA.
XX
AC ABB81792;
XX
DT 24-SEP-2002 (first entry)
XX
DE Human sulfotransferase #1.
XX
KW Human; sulfotransferase; cytostatic; gene therapy; cancer; COPD;
chronic obstructive pulmonary disease.
XX
OS Homo sapiens.
XX
PN WO200253713-A2.
XX
PD 11-JUL-2002.
XX
PF 21-DEC-2001; 2001WO-EP015233.
XX
PR 02-JAN-2001; 2001US-0258812P.
XX
PA (FARB ) BAYER AG.
XX
PI Koehler RH;
XX
WI; 2002-528859/56.
XX
N-PSDB; ABB83168, ABB83170.
XX
Novel human sulfotransferase polypeptide useful in therapeutic methods
for treating disorders such as cancer or chronic obstructive pulmonary
disease, and for screening for human sulfotransferase modulators.
XX
Claim 25; Fig 2; 103pp; English.
XX
The invention relates to a novel human sulfotransferase polypeptide. The
sulfotransferase has cytostatic activity. The polynucleotide of the
invention may have a use in gene therapy. The polypeptide and
polynucleotide of the invention are useful in preventing, ameliorating,
or correcting diseases including cancer and chronic obstructive pulmonary
disease (COPD). The sequence represents a human sulfotransferase of the
invention
XX
Sequence 304 AA;
Alignment Scores:
Pred. No.: 1.45e-30 Length: 304
Score: 422.00 Matches: 97
Percent Similarity: 52.0% Conservative: 59
Best Local Similarity: 32.3% Mismatches: 124
Query Match: 9.6% Indels: 20
Gaps: 5
DB:
US-10-768-158-1 (1-2419) x ABB81792 (1-304)
Qy 21 ATGCGGAGAGCGAGCGGCGAGACCCAGACACCCCGGGGAGTTCGAG---AGCAAGTAC 77
Db 1 MetAlaLysIleGluLysAsnAlaProThrMetGluLysLysProGluLeuPheAsnIle 20
Qy 78 TTCAGTTCATGGGTGGCGTGGCGCCCTTTCGCGCGGGAAGATGAGGAGATCGCC 137
Db 21 MetGluValAspGlyValProThrLeuLeuSerLysGluTrpTrpGluLysValCys 40
Qy 138 AACTTCCCGTGGCGCCAGCGAGTGGGATCGTCACTACCCCAAGTCCGCGCACCGC 197
Db 41 AsnPheGlnAlaLysProAspAspLeuAlaThrTrpProLysSerGlyThrThr 60
Qy 198 TTGCTGCAGAGGTGGTCTACTTGGTGAGCCAGCGCGCTGACCCCGATGAGATCGGCTTG 257
Db 61 TrpMetHisGluIleLeuAspMetIleLeuAsnAspGlyAspValGluLysCysLysArg 80
Qy 258 ATGAACATCGACGAGACTCCCGGTCTCGAG-----TACCCA-----CAGCG 302
Db 81 AlaGlnThrLeuAspArgHisAlaPheLeuGluLeuLysPheProHisLysGluLysPro 100
Qy 303 GGCGTGGACATCATCAAGGAATGACCTCCCGCTCATCAAGAGCCACTGCCCTAC 362
Db 101 AspLeuGluPheValLeuGluMetSerSerProGlnLeuIleLysThrHisLeuProSer 120
Qy 363 CGCTTTCTGCCTCTGACCTCCACAATGGAGACTCCCAAGGTCACTATATGGCTCGCAAC 422
Db 121 HisLeuLeuProProSerIleTrpLysGluAsnCysIleValTrpValAlaArgAsn 140
Qy 423 CCCAAGATCTGGTGGTCTTATATAGTTCCACCGCTCTCTCGGACCATAGGTAC 482
Db 141 ProlysAspCysLeuValSerTy-Tyr-HisPheHisArgMetAlaSerPheMetProAsp 160
Qy 483 CGAGGCACCTTCAAGATCTCGCGGAGTTTATGATGATAGCTGGGTACGGTCC 542
Db 161 ProGlnAsnLeuGluGluPheTy-TyrGluLysPheMetSerGlyLysValValGlyLys 180
Qy 543 TGGTTTGAGCAGTGCAGGAGTCTCGGAGCACCAGCATGGACTCGAAGCTGCTTTTCTC 602
Db 181 TrpPheAspHisValLysGlyTrpTrpAlaAlaLysAspMetHisArgIleLeuTyrl 200
Qy 603 AAGTATGAAGACATGCATCGGACCTGCTGACGATGTTGGAGCAGCTGGCCAGATCTCTG 662
Db 201 PheTyrGluAspIleLysLysAspProlysArgGluIleGluLysIleLeuLysPheLeu 220
Qy 563 GGGGTGCTCTGTGACAGGCCACCTGGAAGCCCTGACGAGCAGTCCACCATCGGTG 722
Db 221 GluLysAspIleSerGluGluIleLeuAsnLysIleIleTyHisThrSerPheAspVal 240
Qy 723 -----GACCATGTGTCAGACGCTGAGGCCCTGCGCGTGGGC----- 758
Db 241 MetLysGlnAsnProMetThrAsnTy-TyrThrLeuProThrSerIleMetAspHisSer 260
Qy 759 -----CGGGGAAGAGTGTGGGTGGAGGACATCTTCACCGTCTCC 800
Db 261 IleSerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTy-TyrPheThrValAla 280
Qy 801 ATGAATGAGAAGTTGACTTGGTGTATATACAGAAAGATGGAAAGTGTGACCTCAGCTTT 860
Db 281 GlnAsnGluGluPheAspLysAspTy-TyrGlnLysLysMetAlaGlySerThrLeuThrPhe 300
```

Db	121	HisLeuIleProProSerIleTrpLysGluAsnCysLysIleValTyrValAlaArgAsn	14
Qy	423	CCCAAGGATCGTGGTCTCTATTATCAGTTCCACCGCTCTCTGGGACCATGAGCTAC	482
Db	141	ProLysAspCysLeuValSerTyrHisPheHisargMetAlaSerPheMetProAsp	160
Qy	483	CGAGCACCCTTTCAGAAATTCGCGGAGGTTTATGAATGATTAAGCTGGGCTACGGCTCC	542
Db	161	ProGlnAsnLeuGluGluPheTyrGluLysPheMetSerGlyLysValValGlyGlySer	180
Qy	543	TGGTTTGACAGCTGCAGAGATTCTGGGAGCACCAGCATGACTGCAACGTGCTTTTCTC	602
Db	181	TrpPheAspHisValIysGlyTrpTrpAlaAlaLysAspMethHisArgIleLeuTyrLeu	200
Qy	603	AAAGTATGAAGACATCATCGGACCTGCTGACGATCGTGAGCAGCTGGCGAGATTCCTG	662
Db	201	PheTyrGluAspIleLysIleAspProLysArgGluIleGluLysIleLeuLysPheLeu	220
Qy	663	GGGGTGTCTGTGACAAAGCCCACTGGAACCCCTGACGGAGCAGTCCACACAGCTGGTG	722
Db	221	GluLysAspIleSerGluGluIleLeuAsnLysIleIleTyrHisThrSerPheAspVal	240
Qy	723	-----GACCAGTGTGCAACGCTGAGCCCTGCGCTGGGGC-----	758
Db	241	MetLysGlnAsnProMetThrAsnTyrThrThrLeuProThrSerIleMetAspHisSer	260
Qy	759	-----CGGGGAAGAGTCTGGGCTGTGGGAAGGACATCTTCACCGTCTCC	800
Db	261	IleSerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAla	280
Qy	801	ATGAATGAGAAGTTTGACTTGGTGTATAAACAGAAAGATGGGAAGTGTGACCTCACGTTT	860
Db	281	GlnAsnGluGluPheAspLysAspTyrGlnLysLysMetAlaGlySerThrLeuThrPhe	300
RESULT 63			
ADD4894	ID	ADD48894 standard; protein; 295 AA.	
XX	AC	ADD48894;	
XX	AC	ADD48894;	
DT	02-DEC-2004	(revised)	
DT	29-JAN-2004	(first entry)	
XX	XX		
DE	Rat Protein P52844, SEQ ID NO 14605.		
XX	Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;		
KW	chronic constriction injury; CCI; spared nerve injury; SNI; Chung.		
XX	Rattus norvegicus.		
OS	Unidentified.		
XX	WQ0203016475-A2.		
PN	27-FEB-2003.		
XX	14-AUG-2002; 2002WO-US025765.		
PF	14-AUG-2001; 2001US-0312147P.		
XX	01-NOV-2001; 2001US-0346382P.		
PR	26-NOV-2001; 2001US-0333347P.		
XX	(GENO) GEN HOSPITAL CORP.		
PA	(FARB) BAYER AG.		
XX	Woolf C, D'urso D, Belfort K, Costigan M;		
XX	WPI; 2003-268312/26.		
DR	GENBANK; P52844.		
XX	New composition comprising two or more isolated polypeptides, useful for		
PT	preparing a medicament for treating pain in an animal.		
XX	Example 1; Page; 1017pp; English.		
PS			

Qy	528	CTGGGCTACGGCTCTCGTTTGAGCAGCGTCGACGAGTTCCTGGGAGCACCAGCATGGACTCG	587
		::: ::::: :::	
Db	167	ValProTyrGlySerTrpTyrAspHisValLysSerTrpTyrGluLysSerLysAsnSer	186
Qy	588	AACGTGCTTTTCTCAAGTATGACACATGCATCGGACCTGGTGACGATGCTGGAGCAG	647
Db	187	ArgValLeuPheMetPheTyrGluAspMetLysGluAspIleArgArgGluValValLys	206
Qy	648	CTGGCCAGATTCTCTGGGGGTGTCTGTGCACAAAGGCCACGCTGGAAGCCCTGACGGAGCAG	707
Db	207	LeuIleGluPheLeuGluArgAspProSerAlaGluLeuValAspArgIleIleGlnHis	226
Qy	708	TGC-----CACCACCTGGTGACACAGTCTGCAAGCTGAGGCCCTGCC-----	752
Db	227	ThrSerPheGlnGluMetLysAsnAsnProCysThrAsnTyrSerMetLeuProGluThr	246
Qy	753	-----GTGGCCGGGGAAGAGTTGGGCTGTGGAAGGAC	785
Db	247	MetIleAspLeuLysValSerProPheMetArgLysGlyIleValGlyAspTrpArgAsn	266
Qy	786	ATCTTACCCTCTCATGAATGAGAAGTTTGACTTGGTGTTATAACAGAGAAGATGGGAAG	845
Db	267	HisPheProGluAlaLeuArgGluArgPheGluGluHisTyrGlnArgHisMetLysAsp	286
Qy	846	TGTGACCTCACGTTT	860
Db	287	CysProValLysPhe	291
RESULT	64		
ABB81793	ID	ABB81793 standard; protein; 304 AA.	
XX	AC	ABB81793;	
XX	DT	24-SEP-2002 (first entry)	
XX	DE	Human sulfotransferase #3.	
XX	KW	Human; sulfotransferase; cytostatic; gene therapy; cancer; COPD;	
XX	KW	chronic obstructive pulmonary disease.	
XX	OS	Homo sapiens.	
XX	PN	WO200253713-A2.	
XX	PD	11-JUL-2002.	
XX	PF	21-DEC-2001; 2001WO-EP015233.	
XX	PR	02-JAN-2001; 2001US-0258812P.	
XX	PA	(FARB) BAYER AG.	
XX	XX	Koehler RH;	
PI	PI	WPI; 2002-528859/56.	
DR	DR		
XX	PT	Novel human sulfotransferase polypeptide useful in therapeutic methods	
PT	PT	for treating disorders such as cancer or chronic obstructive pulmonary	
PT	PT	disease, and for screening for human sulfotransferase modulators.	
XX	PS	Disclosure; Fig 3; 103pp; English.	
XX	CC	The invention relates to a novel human sulfotransferase polypeptide The	
CC	CC	sulfotransferase has cytostatic activity. The polynucleotide of the	
CC	CC	invention may have a use in gene therapy. The polypeptide and	
CC	CC	polynucleotide of the invention are useful in preventing, ameliorating,	
CC	CC	or correcting diseases including cancer and chronic obstructive pulmonary	
CC	CC	disease (COPD). The sequence represents a human sulfotransferase of the	
XX	SQ	Sequence 304 AA;	

Alignment Scores:			
Pred. No.:	2.5e-30	Length:	304
Score:	419.50	Matches:	91
Percent Similarity:	53.0%	Conservative:	57
Best Local Similarity:	32.6%	Mismatches:	112
Query Match:	9.5%	Indels:	19
DB:	5	Gaps:	4
US-10-768-158-1 (1-2419) x ABB81793 (1-304)			
QY	81	GAGTTCATCGCGTGGCGCTGCCGCCCTCTCGCCGGGAAGATGGAGAGATGCCAC	140
DB	22	GlulValAsnGlyIleLeuMetSerLysLeuMetSerAspAsnTrpAspLysIleTrpAsn	41
QY	141	TTCCCGGTGGCGCCAGCGAGCTGTGGATCGTCACCTACCCCAAGTCGGCGACCAAGCTTG	200
DB	42	PheGlnAlaLysProAspAspLeuLeuIleAlaThrTyrAlaLysAlaGlyThrThrTrp	61
QY	201	CTGCAGGAGTGGTCTACTTGGTGTGACCGCGCTGACCCCGATGAGATCGGCTTGATG	260
DB	62	ThrGlnGluIleValAspMetIleGlnAsnAspGlyAspValGlnLysCysGlnArgAla	81
QY	261	AACATCGACGAGCAGCTCCCGTCTGTGAGTAC-----CCACAGCCG-----GCG	305
DB	82	AsnThrTyrAspArgHisProPheIleGluTrpThrLeuProSerProLeuAsnSerGly	101
QY	306	CTGCACATCATCAAGGAACCTCACCTCTCCCGCCCTCATCAAGAGCCACTCGCCCTACCGC	365
DB	102	LeuAspLeuAlaAsnLysMetProSerProArgThrLeuLysThrHisLeuProValHis	121
QY	366	TTTCTGCGCTCTGACCTCCACAATGGAGACTCCAAGGTCACTATATGGCTCGCAACCCC	425
DB	122	MetLeuProProSerPheTrpLysGluAsnSerLysIleIleTyrValAlaArgAsnAla	141
QY	426	AAGATCTGGTGGTCTTATTATCAGTCTCACCGCTCTCTCGCGACCATGAGCTACCGA	485
DB	142	LysAspCysLeuValSerTyrTyrPheSerArgMetAsnLysMetLeuProAspPro	161
QY	486	GGCACCTTTCAAGAATTCTGCCGAGGTTTATGAATGATAGCTGGGTACGGCTCTCGG	545
DB	162	GlyThrLeuGlyGluTyrIleGluGlnPheLysAlaGlyLysValLeuTrpGlySerTrp	181
QY	546	TTTGAGCAGTCGTCAGGAGTCTGGGAGCACCCGATGGACTCGAACCTGCTTTTCTCAAG	605
DB	182	TyrAspHisValLysGlyTrpTipAspValLysAspGlnHisArgIleLeuTyrLeuPhe	201
QY	606	TATGAACACATGCATCGGACCTCGTGACGATGGTGGAGCAGCTGGCCAGATTCCTGGG	665
DB	202	TyrGluAspMetLysGluAspProLysArgGluIleLysIleAlaLysPheLeuGlu	221
QY	666	GTGCTCTGTACAAAGGCCAGCTCGAAGCCCTGACGGAGCACTGCCACAGCTGTG---	722
DB	222	LysAspIleSerGluGluValLeuAsnLysIleIleTyrHisThrSerPheAspValMet	241
QY	723	-----GACCAGTCTGCGAACCGCTGAGGCCCTCGCCC-----752	
DB	242	LysGluAsnProMetAlaAsnTyrThrThrLeuProSerSerIleMetAspHisSerIle	261
QY	753	-----GTGGGCGGGGAAGAGTGGGCTGTGGGAAGGACATCTTCACGCTCTCCATG	803
DB	262	SerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAlaGln	281
QY	804	AATCAGAAGTTTGACTTGGTGTATAAACAGAGAAGTGGAAAGTGTGACCTCACGTTT	860
DB	282	SerGluAspPheAspGluAspTyrArgArgLysMetAlaGlySerAsnIleThrPhe	300
RESULT 65			
ID	AD117047		
XX	AD117047 standard; protein: 304 AA.		
AC	AD117047;		
XX	AD117047;		
DT	15-APR-2004 (first entry)		


```
QY 366 TTTCTGCCCTCTGACCTCCACATCGAGACTCCAAAGGTCACTATATATGGCTCGCAACCCC 425
D 122 MetLeuProProSerPheTrpIysGluAsnSerIysIleIleTyrValAlaArgAla 141
QY 456 AAGGATCTGGTGGTCTTATTATCAGTTTCCACCGCTCTCTCGCGACCATGAGCTACCGA 485
D 142 LysAspCysLeuValSerIleTyrPheSerArgMetAsnLysMetLeuProAspPro 161
QY 486 GGCACCTTTCAAGAAATTCGCGGAGGTTTATGAATGATGAAGCTGGCTCGGCTCTCTGG 545
D 162 GlyThrLeuGlyGluTyrIleGluGlnPheLysAlaGlyLysValLeuTrpGlySerTrp 181
QY 546 TTGAGCAGCTGCAGGAGTTCTGGAGACCGCATGGACTCGAAGCTGCTTTTCTCAAG 605
D 182 TyrAspHisValLysGlyTrpTrpAspValLysAspGlnHisArgIleLeuTyrLeuPhe 201
QY 606 TATGAAGACATGCATCGGAGCTGTGTGACGATGGTGGAGCAGCTGGCCAGATCTCTGGGG 665
D 202 TyrGluAspMetLysGluAspProLysArgGluIleLysIleAlaLysPheLeuGlu 221
QY 666 GTGTCTCTGCAAGCCGAGCTGGAAGCCCTGACGAGCAGCTGCCACCATGCTGTG--- 722
D 222 LysAspIleSerGluGluValLeuAsnLysIleIleTyrHisThrSerPheAspValMet 241
QY 723 -----GACCAGTGTGCAAGCTGAGGCCCTGCCCC----- 752
D 242 LysGluAsnProMetAlaAsnTyrThrThrLeuProSerSerIleMetAspHisSerIle 261
QY 753 -----GTGGCGCGGGAAGAGTGTGGGCTGTGGAAGGACATCTTCACCGTCTCCATG 803
D 262 SerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAlaGln 281
QY 804 AATGAGAAGTTTGACTGCTGTGTATAAACAAGATGGGAAAGTGTGACCTCACCGTTT 860
D 282 SerGluAspPheAspGluAspTyrArgArgLysMetAlaGlySerAsnIleThrPhe 300

RESULT 66
ABB81794
ID ABB81794 standard; protein; 283 AA.
AC ABB81794;
DX 24-SEP-2002 (first entry)
DE Human sulfotransferase #2.
KW Human; sulfotransferase; cytostatic; gene therapy; cancer; COPD;
KW chronic obstructive pulmonary disease.
OS Homo sapiens.
XX WO200253713-A2.
XX 11-JUL-2002.
XX 21-DEC-2001; 2001WO-EP015233.
XX 02-JAN-2001; 2001US-0258812P.
XX (FARB ) BAYER AG.
XX Koehler RH;
XX WPI; 2002-528859/56.
XX N-PSDB; ABB83171.
XX Novel human sulfotransferase polypeptide useful in therapeutic methods
XX for treating disorders such as cancer or chronic obstructive pulmonary
XX disease, and for screening for human sulfotransferase modulators.
XX Disclosure; Fig 6; 103pp; English.
```

CC The invention relates to a novel human sulfotransferase polypeptide The
CC sulfotransferase has cytostatic activity. The polynucleotide of the
CC invention may have a use in gene therapy. The polypeptide and
CC polynucleotide of the invention are useful in preventing, ameliorating,
CC or correcting diseases including cancer and chronic obstructive pulmonary
CC disease (COPD). The sequence represents a human sulfotransferase of the
CC invention

XX SQ Sequence 283 AA;

Alignment Scores: Length: 283
Pred. No.: 3e-30 Matches: 89
Score: 418.50
Percent Similarity: 54.5% Conservative: 55
Best Local Similarity: 33.7% Mismatches: 101
Query Match: 9.5% Indels: 19
DB: Gaps: 4

US-10-768-158-1 (1-2419) x ABB81794 (1-283)

```
QY 126 GAGGAGATCCGCCAACTTCCCGGTGCGGCCAGCAGCAGCTGTGGATCGTCACCTACCCCAAG 185
D 16 GluLysValCysAsnPheGlnAlaLysProAspAspLeuIleLeuAlaThrTyrProLys 35
QY 186 TCCGGCACCGACTTGTCTGCGAGGAGTGTCTACTTGTGTAGCCAGGGCGGTGACCCCGAT 245
D 36 SerGlyThrTrpMetHisGluIleLeuAspMetIleLeuAsnAspGlyAspValGlu 55
QY 246 GAGATCGGCTTGATGAACATCGACGACGAGCTCCCGTCTCTGGAG-----TACCCA--- 296
D 56 LysCysLysArgAlaGlnThrLeuAspArgHisAlaPheLeuGluLeuLysPheProHis 75
QY 297 -----CAGCGCGGCGCTGGACATCAAGAACTCTCTCCCGCTCATCAAGAGC 350
D 76 LysGluLysProAspLeuGluPheValLeuGluMetSerSerProGlnLeuIleLysThr 95
QY 351 CACTGCGCTACCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 410
D 96 HisLeuProSerHisLeuIleProSerIleTrpLysGluAsnCysLysIleValTyr 115
QY 411 ATGGCTCGCAACCCCAAGGATCTGGTGTGCTTATTATCAGTTTCCACCGCTCTCTGCGG 470
D 116 ValAlaArgAsnProLysAspCysLeuValSerTyrTyrHisPheHisArgMetAlaSer 135
QY 471 ACCATGAGTACCGGAGCACCTTTCAAGAATTTCTGCCGAGGTTTATGAATGATAAGCTG 530
D 136 PheMetProAspProGlnAsnLeuGluGluPheTyrGluLysPheMetSerGlyLysVal 155
QY 531 GGCTACCGCTCTCGTGTGGAGCAGCTGAGGAGTTCTGGAGAGCACCGCATGGATCGAAC 590
D 156 ValGlyGlySerTrpPheAspHisValLysGlyTrpTrpAlaAlaLysAspMetHisArg 175
QY 591 GTGCTTTTCTCAAGTATGAACATCATCGGAGCCTGCTGACGATGGTGGAGCAGCTG 650
D 176 IleLeuTyrLeuPheTyrGluAspIleLysLysAspProLysArgGluIleGluLysIle 195
QY 651 GCCAGATTCTCGGGGTGCTCTGTGACAGAGCCCGAGCTGGAAAGCCCTGAGGAGCAGCTGC 710
D 196 LeuLysPheLeuGluLysAspIleSerGluGluIleLeuAsnLysIleIleTyrHisThr 215
QY 711 CACGAGCTGTG-----GACCAGTGTGCAACGCTGAGGCCCTGCGCGTGGGC--- 758
D 216 SerPheAspValMetLysGlnAsnProMetThrAsnTyrThrThrLeuProThrSerIle 235
QY 759 -----CGGGGAAGAGTTGGGCTGTGGAGGAGCATC 788
D 236 MetAspHisSerIleSerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyr 255
QY 789 TTCAACGCTCTCCATGAATGAGAAGTTTGACTTGTGTATATAACAGAGATGGGAAAGTGT 848
D 256 PheThrValAlaGlnAsnGluGluPheAspLysAspTyrGlnLysLysMetAlaGlySer 275
QY 849 GACCTCACGTTT 860
```

Db	276 ThrLeuThrPhe 279		Query Match: 9.3%	Indels: 19
AAAM50843			5	3
AAAM50843	standard; protein; 304 AA.			
AC	AAAM50843;			
DT	01-MAY-2002 (first entry)			
DE	Orphan G protein-coupled receptor GPR-22.			
XX	Orphan G protein-coupled receptor; GPR-22; mouse; transgenic mouse;			
KW	transgenic animal; animal model; gene disruption; gene targeting;			
KW	gene detection; gene therapy; receptor.			
XX				
OS	Mus musculus.			
XX				
PN	WO200206445-A2.			
XX				
PD	24-JAN-2002.			
XX				
PF	17-JUL-2001; 2001WO-US022402.			
XX				
PR	19-JUL-2000; 2000US-0219168P.			
PR	19-JUL-2000; 2000US-0219171P.			
PR	19-JUL-2000; 2000US-0219175P.			
PR	25-JUL-2000; 2000US-0221455P.			
PR	13-DEC-2000; 2000US-0256212P.			
XX				
PA	(DELT-) DELTAGEN INC.			
XX				
PI	Brennan TV, Allen KD;			
XX				
DR	WPI; 2002-179787/23.			
DR	N-PSDB; ABA91712.			
XX				
PT	Non-human transgenic animal useful as a disease model and for identifying			
PT	agents that modulate gene expression and function, comprises a disruption			
PT	in a targeted gene e.g. cysteine protease-like gene.			
PS	Example 4; Fig 7; 61pp; English.			
XX				
CC	The present sequence is that of the protein encoded by the mouse orphan G			
CC	protein-coupled receptor GPR-22 gene (see ABA917120). The invention			
CC	provides transgenic animals containing targeted gene disruptions,			
CC	including disruptions of the GPR-22 gene. A claimed gene targeting			
CC	construct comprises a first polynucleotide sequence homologous to a			
CC	target gene, a second polynucleotide sequence homologous to the target			
CC	gene, a selectable marker and optionally a screening marker. The			
CC	targeting construct is obtained by inserting a first and a second			
CC	polynucleotide homologous to different regions of the target gene into a			
CC	vector so that they flank a positive selection marker. A cell, especially			
CC	a murine embryonic stem cell, and a transgenic animal comprising a			
CC	disruption in a target gene are claimed. In the present case, a GPR-22-			
CC	specific targeting construct having the ability to disrupt or modify GPR-			
CC	22 genes was created using the targeting arms (homologous sequences)			
CC	given in ABA91713 and ABA91714. Transgenic mice were generated. The cell-			
CC	and animal-based systems are useful as models for disease and for			
CC	identifying agents that modulate gene expression and function, and as			
CC	potential treatments for various disease states and disease conditions.			
CC	Methods of treating diseases associated with disrupted targeted gene			
CC	expression or function comprise detecting and replacing mutated target			
XX	genes through gene therapy			
SQ	Sequence 304 AA;			
Alignment Scores:				
Pred. No.:	2.8e-29	Length:	304	
Score:	408.50	Matches:	88	
Percent Similarity:	52.0%	Conservative:	57	
Best Local Similarity:	31.5%	Mismatches:	115	

US-10-768-158-1 (1-2419) x AAM50843 (1-304)				
Qy	81	GAGTTCCATGCGGTGCGGCTGCGCCCTTCTGCGCGGGAAGATGGAGAGATCGCCAAC	140	
Db	22	GluValAsnGlyIleLeuMetSerLysMetMetSerGluAsnTrpAspLysIleTrpAsn	41	
Qy	141	TTCCCGGTGCGGCCACGACGTGTGGATCGTCACCTACCCCAAGTCCGCGCACCGTTG	200	
Db	42	PheGlnAlaLysProAspAspLeuLeuIleAlaThrTyrAlaLysAlaGlyThrTrp	61	
Qy	201	CTGCAGAGGTGCTACTTGGTCAGCGCGCTGACCCCGATGAGATCGGCTTGTATG	260	
Db	62	ThrGlnGluIleValAspMetIleGlnAsnAspGlyAspValGlnLysCysGlnArgAla	81	
Qy	261	AACATCGACGACGACCTCCCGCTCTGGAGTACCCACACGCGC-----GGC	305	
Db	82	AsnThrTyrAspArgHisProPheIleGluTrpThrLeuProProProLeuAsnSerGly	101	
Qy	306	CTGGACATCATCAAGAACTGACCTCTCCCGCTCTATCAAGACCACTGCTCCCTACCGC	365	
Db	102	LeuAspLeuAlaAsnLysMetProSerProArgThrLeuLysThrHisLeuProValGln	121	
Qy	366	TTTCTGCTCTGACCTCCCAATGGAGCTCCCAAGCTCATATATGGTCTCGCAACCCC	425	
Db	122	MetLeuProProSerPheTrpLysGluAsnSerGlnIleIleIleIleIleIleIleIle	141	
Qy	426	AAGGATCTGTGTGTCTTATTATTCAGTTCTCCACCGCTCTCTGCGGACCATGAGCTAC	485	
Db	142	LysAspCysLeuValSerTyrTyrPheSerArgMetAsnLysMetLeuProAspPro	161	
Qy	486	GGACCTTTCAAGAAATCTGCGGAGGTTTATGATGATTAAGCTGGGCTCGGCTCTGG	545	
Db	162	GlyThrLeuGlyGluTyrIleGluThrPheLysAlaGlyLysValLeuTrpGlySerTrp	181	
Qy	546	TTTGAGCAGCTGCGAGGAGTTCTGGGACACCGCATGCTCAACGCTCTTTCTCAAG	605	
Db	182	TyrAspHisValLysGlyTrpTrpAspValLysAspLysHisArgIleLeuTrpPhe	201	
Qy	606	TATGAGACATGTCATCGGACCTGGTGACGATGGTGAGCAGCTGGCCAGATTCTCGGG	665	
Db	202	TyrGluAspMetLysGluAspProLysArgGluIleLysLysIleValLysPheLeuGlu	221	
Qy	666	GTGTCCTGTGACAGGCCCGACCTGGAAGCCCTGACGAGCAGCTCCACCGAGCTGGT	722	
Db	222	LysAspIleSerGluGluValLeuAsnLysIleIleHisThrSerPheAspValMet	241	
Qy	723	-----GACCAGTGTCTCAACGCTGAGCGCTGCGC-----	752	
Db	242	LysGlnAsnProMetAlaAsnTyrThrLeuProSerSerIleMetAspHisSerIle	261	
Qy	753	-----GTGGCCCGGGAAGTGGGTGTGGGAAGGACATCTTCCCGTCTCCATG	803	
Db	262	SerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAlaGln	281	
Qy	804	AATCAGAGTTTGAATGGTGTATTAACAGAGATGGAAAGTGTGACCTCACGTTT	860	
Db	282	SerGluAspPheAspGluAspTyrArgLysLysMetAlaGlySerThrIleThrPhe	300	
RESULT 68				
AD117048				
ID	AD117048 standard; protein; 304 AA.			
XX	AC			
XX	AD117048;			
XX				
DT	15-APR-2004 (first entry)			
XX				
DE	Murine NOVX protein homologue SeqID 584.			
XX				
KW	mouse; murine; NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes;			
KW	inflammation; autoimmune disorder; allergy; blood disorder;			

acquired immunodeficiency syndrome; AIDS; obesity; asthma;
immunoglobulin (Ig)A nephropathy; cirrhosis; arthritis;
Alzheimer's disease; infect.

Mus musculus.

WO200268649-A2.

06-SEP-2002.

31-JAN-2002; 2002WO-US002785.

31-JAN-2001; 2001US-0265339P.

31-JAN-2001; 2001US-0265412P.

31-JAN-2001; 2001US-0265514P.

02-FEB-2001; 2001US-0265517P.

05-FEB-2001; 2001US-0266406P.

07-FEB-2001; 2001US-0266767P.

07-FEB-2001; 2001US-0266975P.

08-FEB-2001; 2001US-0267459P.

09-FEB-2001; 2001US-0267823P.

15-FEB-2001; 2001US-0268974P.

26-FEB-2001; 2001US-0271664P.

27-FEB-2001; 2001US-0271833P.

27-FEB-2001; 2001US-0271855P.

02-MAR-2001; 2001US-0272788P.

02-MAR-2001; 2001US-0275989P.

15-MAR-2001; 2001US-0276448P.

15-MAR-2001; 2001US-0276450P.

16-MAR-2001; 2001US-0276397P.

16-MAR-2001; 2001US-0276768P.

20-MAR-2001; 2001US-0278652P.

26-MAR-2001; 2001US-0278775P.

26-MAR-2001; 2001US-0278778P.

29-MAR-2001; 2001US-0279882P.

29-MAR-2001; 2001US-0279884P.

30-MAR-2001; 2001US-0280147P.

11-APR-2001; 2001US-0282992P.

11-APR-2001; 2001US-0283083P.

20-APR-2001; 2001US-0285133P.

23-APR-2001; 2001US-0285749P.

03-MAY-2001; 2001US-0288327P.

03-MAY-2001; 2001US-0288504P.

29-MAY-2001; 2001US-0294047P.

30-MAY-2001; 2001US-0294473P.

08-JUN-2001; 2001US-0296964P.

18-JUN-2001; 2001US-0298959P.

KW

KW

KW

XX

XX

OS

XX

XX

PN

XX

PD

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

```
Qy 426 AAGGATCGGTGGTGTCTATTATCAGTTCCACCGCTCTCTCGCGACCATGAGCTACCGA 485
    ||||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 142 LysAspCysLeuValSerTyrTyrPheSerArgMetAsnLysMetLeuProAspPro 161
    ||||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 486 GGCACCTTTCAGAAATCTGCGGAGGTTTATGAATGATAAGCTGGGCTACGGCTCTCTGG 545
    ||||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 162 GlyThrLeuGlyGluTyrIleGluThrPheLysAlaGlyLysValLeuTyrGlySerTrp 181
    ||||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 546 TTTGAGCAGCTGCAGGAGTTCTGGGAGCACCGCATGGACTCGAAGCTGCTTTTCTCAAG 605
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 182 TyrAspHisValLysGlyTyrTrpAspValLysAspLysHisArgIleLeuTyrLeuPhe 201
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 606 TATGAAGACATGCATCGGACCTGCTGACGATGGTGGAGCAGCTGGCCGAGATTCTCTGGG 665
    ||||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 202 TyrGluAspMetLysGluAspProLysArgGluLeuLysLysIleValLysPheLeuGlu 221
    ||||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 666 GTGCTCTGTGAAGCCACCTGGAAGCCCTGAGCGGAGCAGCTGCCACCACTGGTG--- 722
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 222 LysAspIleSerGluGluValLeuAsnLysIleIleHisHisThrSerPheAspValMet 241
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 723 -----GACCAGTGTGCAAGCTGAGGCCCTGCCC----- 752
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 242 LysGlnAsnProMetAlaAsnTyrThrThrLeuProSerSerIleMetAspHisSerIle 261
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 753 -----GTGGCGGGGAAGAGTTGGGCTGTGGAGGACATCTTCACCGTCTCCATG 803
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 262 SerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAlaGln 281
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 804 AATGAGAAGTTTGACTGGTGTATATAACAGAAAGATGGGAAGTGTCACCTCACGTTT 860
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 282 SerGluAspPheAspGluAspTyrArgLysLysMetAlaGlySerThrIleThrPhe 300

RESULT 69
ID AAE12210 standard; protein; 304 AA.
XX
AC AAE12210;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human ST drug-metabolising protein 3 encoded by cDNA.
XX
KW Human; sulfotransferase; drug-metabolising enzyme mediated disorder; ST;
XX DME; toxicity study; gene therapy.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 173..175
FT /note="Protein kinase C phosphorylation site"
FT Modified-site 180..183
FT /note="Casein kinase II phosphorylation site"
FT Modified-site 194..196
FT /note="Protein kinase C phosphorylation site"
FT Modified-site 236..239
FT /note="Casein kinase II phosphorylation site"
FT Modified-site 248..251
FT /note="Asn is N-glycosylated"
FT Modified-site 258..261
FT /note="Asn is N-glycosylated"
FT Modified-site 295..300
FT /note="N-myristoylation site"
XX
FN WO200172977-A2.
XX
PD 04-OCT-2001.
XX
PP 27-MAR-2001; 2001WO-US009659.
XX
PR 27-MAR-2000; 2000US-0192408P.
XX 20-JUN-2000; 2000US-0212725P.
XX 03-JUL-2000; 2000US-00609816.
XX
```

```
PA (PEKE ) PE CORP NY.
XX
PI Woodage T, Wei M, Kodira C, Di Francesco V, Beasley EM;
XX
WPI; 2001-616484/71.
DR N-PSDB; AAD19778.
XX
PT New polypeptides useful in labeling reagents with tagged sulfur and
FT identifying modulators of the protein comprises proteins related to
FT sulfotransferase drug-metabolizing enzyme subfamily.
XX
PS Claim 1; Fig 2; 83pp; English.
XX
CC The invention relates to sulfotransferase (ST) drug-metabolising enzyme
CC (DME) peptides, proteins and nucleic acid molecules encoding them.
CC Sequences of the invention are useful in labelling reagents with tagged
CC sulfur, for the development of human therapeutics, as query sequences to
CC perform a search against sequence databases, for identifying other family
CC members or related sequences, to raise antibodies or to elicit another
CC immune response, as a reagent in assays designed to quantitatively
CC determine levels of the protein in biological fluids, to metabolise
CC compounds for toxicity studies, and as markers for tissues in which the
CC corresponding protein is preferentially expressed. Polypeptides of the
CC invention is also useful in diagnosing a disease, predisposition to a
CC disease mediated by the peptide, in pharmacogenomic analysis and treating
CC a disorder characterised by absence of inappropriate or unwanted
CC expression of the protein. Polynucleotides of the invention are useful
CC for constructing recombinant vectors and host cells, expressing antigenic
CC portions of the proteins, in genotyping and mapping experiments, for
CC designing ribozymes, for monitoring the effectiveness of modulating
CC compounds on the expression or activity of the protein, in diagnostic
CC assays, antisense constructs to control protein expression and in gene
CC therapy. The present sequence is ST drug-metabolising protein 3 encoded
CC by cDNA
XX
SQ Sequence 304 AA;
XX
Alignment Scores:
Pred. No.: 3.9e-29 Length: 304
Score: 407.00 Matches: 94
Percent Similarity: 51.7% Conservative: 61
Best Local Similarity: 31.3% Mismatches: 125
Query Match: 9.2% Indels: 20
DB: 4 Gaps: 5
US-10-768-158-1 (1-2419) x AAE12210 (1-304)
Qy 21 ATGCGGAGAGCGGCGGAGACCCCGAGACCCCGGGGAGTTCGAG---AGCAAGTAC 77
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MetAlaLysIleGluLysAsnAlaProThrMetGluLysLysProGluLeuPheAsnIle 20
Qy 78 TTGAGTTCATGCGTGGCGGCTGCCGCCCTTTCGCCGGGAGATGGAGGAGATGCC 137
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 21 MetGluValAspGlyValProThrLeuIleLeuSerLysGluTyrTrpGluLysValCys 40
Qy 138 AACTTCCCGTGGCGGCGGAGCGAGTGTGGATCGTCACTACCCCAAGTCGGGACCCAGC 197
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 41 AsnPheGlnAlaLysProAspAspLeuIleLeuAlaThrTyrProLysSerGlyThrThr 60
Qy 198 TTGCTGCAGGAGTGTGCTACTTGTGTCAGCGCGGCTGACCCCGATGAGATCGGCTTG 257
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 TrpMethisGluIleLeuAspMetIleLeuAsnAspGlyAspValGluLysCysLysArg 80
Qy 258 ATGAACATCGACGAGCAGCTCCCGGCTCTCGAG-----TACCCA-----CAGCG 302
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 81 AlaGlnThrLeuAspArgHisAlaPheLeuLysPheProHisLysGluLysPro 100
Qy 303 GGCTTGACATCATCAAGGAACTGACCTCTCCCGCTCATCAAGAGCCACCTCCCTAC 362
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 101 AspLeuGluPheValLeuGluMetSerSerProGlnLeuLysThrHisLeuProSer 120
Qy 363 CGCTTTCTGCCCTCTGACCTCCCAATGGAGACTCCCAAGTGCATCTTATATGGTTCGCAAC 422
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

Db 121 HisleulleProProSerIleTrpLysGluAsnCysLysIleValTyrValAlaArgAsn 140
Qy 423 CCCAAGGATCTGGTGTCTTATATATCAGTTTCACCGCTCTCTCGGACCATGAGTAC 482
Db 141 ProlysAspCysLeuValSerTyrHisPheHisArgMetAlaSerPheMetProAsp 160
Qy 483 CGAGGCACCTTCAAGAAATCTCGCGGAGGTTTATGAATGATAGCTGGGTACGGTCC 542
Db 161 ProGlnAsnLeuGluPheTyrGluLysPheMetSerGlyLysValValGlySer 180
Qy 543 TGGTTTGACGCTGCAGAGTCTCGGAGCACCGCATGGACTCGAAGCTCTTTTCTC 602
Db 181 TrpPheAspHisValLysGlyTrpAlaAlaLysAspThrHisArgIleLeuTyrLeu 200
Qy 603 AAGTATGAAGACATCGGACCTGTGTACGATGTGGAGCAGCTGGCCAGATTCCTG 662
Db 201 PheTyrGluAspIleLysAsnProLysHisGluIleHisLysValLeuGluPheLeu 220
Qy 663 GGGGTGTCCTGTACAGGCCAGCTGGAAGCCCTGACGGAGCACTGCCACCACTGGTG 722
Db 221 GlulysThrLeuSerGlyAspValIleAsnLysIleValHisHisThrSerPheAspVal 240
Qy 723 -----GACCAGTGTGCAAGCTGAGGCCCTGCC----- 752
Db 241 MetLysAspAsnProMetAlaAsnHisThrAlaValProAlaHisIlePheAsnHisSer 260
Qy 753 -----GTGGCGCGGGAAGAGTGTGGGCTGTGGAGGACATCTTCACCGCTCTCC 800
Db 261 IleSerLysPheMetArgLysGlyMetProGlyAspTrpLysAsnHisPheThrValAla 280
Qy 801 ATGAATGAGAAGTTGACTTGGTGTATAAACAGAAAGATGGGAAAGTGTGACCTCACGTTT 860
Db 281 MetAsnGluAsnPheAspLysHisTyrGluLysLysMetAlaGlySerThrLeuAsnPhe 300
RESULT 70
ID ABB66006 standard; protein; 316 AA.
XX AC ABB66006;
XX DT 26-MAR-2002 (first entry)
XX DE Drosophila melanogaster polypeptide SEQ ID NO 24810.
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical.
XX OS Drosophila melanogaster.
XX PN WO200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US009231.
XX PR 23-MAR-2000; 2000US-0191637P.
XX PR 11-JUL-2000; 2000US-00614150.
XX PA (PEKE) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX DR N-PSDB; ABL10109.
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signaling and cell-cell
XX PT interactions.
XX PS Disclosure; SEQ ID NO 24810; 21pp + Sequence Listing; English.
XX XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABU16176-ABLJ30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-
CC ABBS72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 316 AA;
Alignment Scores: 6.17e-29 Length: 316
Pred. No.: 94 Matches: 94
Score: 405.00
Percent Similarity: 53.5% Conservative: 52
Best Local Similarity: 34.4% Mismatches: 99
Query Match: 9.2% Indels: 28
DB: 4 Gaps: 8
US-10-768-158-1 (1-2419) x ABB66006 (1-316)
Qy 126 GAGGAGATCGCAACTTCCCGGTGCGGCCCGAGCGAGCTGTGATCGTCACTACCCCAAG 185
Db 45 GluSerIleArgSerLeuProValTyrGlnAspAspValTrpMetValSerTyrProArg 64
Qy 186 TCCGGCACCACTGTGCTGCAGAGGTGTCTACTTGTGTGAGCCAG----- 230
Db 65 ThrGlySerThrTrpAlaGlnGluMetValTrpLeuLeuGlyHisGlnLeuAspTyrVal 84
Qy 231 GGGCTGACCCCGAT-----GAGATC-----GGCTTGATGAAC 263
Db 85 AlaAlaGlnAspLeuArgLeuArgSerProLeuIleGluLeuSerAlaLeuPheSer 104
Qy 264 ATGACGAGCAGCTCCCGGTCTGGAGTACCCACAGCCGGCGCTGGACATCATCAAGAA 323
Db 105 IleAspHisGluThrValAlaGlnLysPheGlyAsnThrValAspLeuValArgAsn 124
Qy 324 CTGACCTCTCCCGCTCATCAAGAGCCACCTGCGCTTACCGCTTTCTTGCCTCTTGACCTC 383
Db 125 LeuProArgProArgPheAlaArgSerHisLeuProTrpProLeuLeuProGluGlnPhe 144
Qy 384 CACATGGAGACTCCAAAGTCTATATATGCTGCGCAACCCCAAGGATCTGTGGTGTCT 443
Db 145 GluThrValLysProArgIleValTyrThrAlaArgAsnProLysAspLeuValSer 164
Qy 444 TATTATCAGTTCCACCGCTCTCTGCGGACCATGAGTACCGAGGCACCTTTCAAGATTC 503
Db 165 TyrTyrHisTyrPheLysLeuLeuHisGlyMetAsn-----GlyAspPheGluGlnPhe 182
Qy 504 TGCCGGAGGTTTATGATGATAAGCTGGGCTACGGCTCTGTGGTGTGAGCAGTGCAGAG 563
Db 183 ValAspLeuPheLeuGluGlyHisThrProMetGlySerTyrTrpArgHisValLeuPro 202
Qy 564 TTCTGGGAGCACCGCATGGACTCGAAGCTCTTTTCTCAAGTATGAGACATGATCGG 623
Db 203 PheTrpLysArgSerGlnAspAspAsnValLeuPheIleLysTyrGluAspMetValLys 222
Qy 624 GACCTGTGACGATGGTGGAGCAGCTGGCCAGATTCCTGGGGGTGTCC-----TGTGAC 677
Db 223 AspLeuProSerValValArgArgCysAlaArgPheLeuGlyValGlnSerLeuLeuAsp 242
Qy 678 AAGGCCAGCTGGAAGCCCTGACGGAGCAGTGC-----CACCAGCTGGTGGACAG 728
Db 243 ValSerThrLeuGlnLysLeuCysAspHisLeuThrPheAspLysMetArgAlaAsnLys 262
Qy 729 TGCTGCAACCTGAG---GCCCTGCC-----GTGGCGCGGGGAAGA 767
Db 263 AlaValAsnLeuGluLysLeuLeuProGluSerSerLysPheIleArgAsnGlyLys 282
Qy 768 GTTGGGCTGTGGAAGGACATCTTCCCGTCTCCATGAATGAGAAAGTTTGACTTGGTGTAT 827
Db 283 IleGlyAspTrpArgAsnHisMetGlyAsnGluMetSerGluArgPheAspGluTrpThr 302

Pred. No.:	2, 03e-28	Length:	305
Score:	399.50	Matches:	93
Percent Similarity:	51.2%	Conservative:	61
Best Local Similarity:	30.9%	Mismatches:	126
Query Match:	9.1%	Indels:	21
DB:	6	Gaps:	5
US-10-768-158-1 (1-2419) x ABU12099 (1-305)			
Qy	21	ATGCGGAGAGCGGCGGAGACCCCGGCGGAGTTCGAG---AGCAAGTAC	77
Db	1	MetAlaLysThrGluLysAsnAlaProThrMetGluLysLysProGluLysPheAsnIle	20
Qy	78	TTCCAGTTCCATGGCGTGGCGTCCGCCCTTCTCCCGCGGAGAGGAGATCC	137
Db	21	MetGluValAspGlyValProThrLeuLeuSerLysGluTrpTrpGluLysValCys	40
Qy	138	AACCTTCCCGTGGCGGCGGAGCGTGTGGATCGTCACCTACCCCAAGTCGGGACACG	197
Db	41	AsnPheGlnAlaLysProAspAspLeuIleAlaThrTyrProLysSerGlyThrThr	60
Qy	198	TTGCTGCGAGAGTGTCTACTTGTGTGAGCCGCGCTGACCCCGATGAGATCGGCTTG	257
Db	61	TrpMetHisGluIleLeuAspMetIleLeuAsnAspGlyAspValGluLysCysLysArg	80
Qy	258	ATGAACATCGAGGACGCTCCGGTCTCGAG-----TACCCA-----CAGCG	302
Db	81	AlaGlnThrLeuAspArgHisAlaPheLeuGluLeuLysPheProHisLysGluLysPro	100
Qy	303	GGCTTGGACATCATCAAGAACTACCTCTCCCGCCTCATCAAGAGCCACTGCCCTAC	362
Db	101	AspLeuGluPheValLeuGluMetSerProGlnLeuIleLysThrHisLeuProSer	120
Qy	363	CGCTTTCGCCCTCTGACCTCCACAAATGGAGACTCCAAAGTCATCTATATGCTCGCAAC	422
Db	121	HisLeuIleProProSerIleTrpLysGluAsnCysLysIleValValAlaArgAsn	140
Qy	423	CCCAAGATCTGGTGTCTTATTATTCAGTTCCACCGCTCTCTGCGGACCATGAGTAC	482
Db	141	ProLysAspCysLeuValSerTyrTyrHisPheHisArgMetAlaSerPheMetProAsp	160
Qy	483	CGAGGCACCTTTCAAGAACTTCGCGGAGGTTTATGAATGATAAGCTGGGCTACGGCTCC	542
Db	161	ProGlnAsnLeuGluGluPheTyrGluLysPheMetSerGlyLysGlyGluPheGlySer	180
Qy	543	TGGTTGACGACGTGCAGGAGTTCTGGGAGCAGCCGATGAGCTCGAAAGTCTTTTCTC	602
Db	181	TrpPheAspHisValLysGlyTrpTrpAlaLysAspMetHisArgIleLeuTyrLeu	200
Qy	603	AAGTATGAACATGCTCGGACCTGGTACGATGGTGGAGCAGCTGGCCAGATTCTCTG	662
Db	201	PheTyrGluAspIleLysGlnAsnProLysHisGluIleHisLysValLeuGluPheLeu	220
Qy	663	GGGTGTCTGTGACAGGCCACGTGGAAGCCCTGACGGAGCAGTCCACCATCGTGGT	722
Db	221	GluLysThrTrpSerGlyAspValIleAsnLysIleValHisHisThrSerPheAspVal	240
Qy	723	-----GACAGTGTGCAAGCTGAGCGCTCGCCCTG-----	755
Db	241	MetLysAspAsnProMetAlaAsnHisThrAlaValProAlaHisIlePheAsnHisSer	260
Qy	756	-----GGCGGGGAGAGTTGGGCTTGGTGGAGGACATCTTCCCGTC	797
Db	261	IleSerLysPheMetArgLysGlyMetProGlyAspTrpLysAsnHisPheThrVal	280
Qy	798	TCCATGATGAGAGTTTACTTGGTGTATAAAGAGAGATGGGAAAGTGTGACCTCAG	857
Db	281	AlaLeuAsnGluAsnPheAspLysHisTyrGluLysLysMetAlaGlySerThrLeuAsn	300
Qy	858	TTT 860	
Db	301	Phe 301	

RESULT 74	
AAE05178	
ID	AAE05178 standard; protein; 304 AA.
AC	AAE05178;
XX	
DT	12-SEP-2001 (first entry)
XX	
DE	Human drug metabolising enzyme (DME-9) protein.
XX	
KW	Human; drug metabolising enzyme; DME-9; immunosuppressive; gene therapy; cytostatic; autoimmune disorder; inflammatory disorder; atherosclerosis; osteoporosis; eye disorder; hepatic tumour; Addison's disease; cretinism; rheumatoid arthritis; acquired immune deficiency syndrome; AIDS; anaemia; developmental disorder; endocrine disorder; iritis; acromegaly; epilepsy; thyrotoxicosis; pancreatic disorder; diabetes mellitus; obesity; adenoma; gastrointestinal disorder; nodular hyperplasia; conjunctivitis; glaucoma; actinic keratosis; metabolic disorder; dysphagia; anorexia; carcinoma; cell proliferative disorder.
XX	
OS	Homo sapiens.
XX	
Key	Location/Qualifiers
Binding-site	20..304
FT	/label= PAPS_cofactor_binding_site.
ET	
XX	WO200151638-A2.
PN	
XX	19-JUL-2001.
PD	
XX	12-JAN-2001; 2001WO-US001174.
PF	
XX	14-JAN-2000; 2000US-0176139P.
PR	21-JAN-2000; 2000US-0177443P.
PR	28-JAN-2000; 2000US-0178574P.
XX	
PA	(INCY-) INCYTE GENOMICS INC.
XX	
XX	Yang J, Baughn MR, Burford N, Au-Young J, Lu DAM, Reddy R;
PI	Ring HZ, Hillman JL, Yue H, Azimzai Y, Yao MG, Gandhi AR;
PI	Nguyen DB, Tang YT, Lal P, Bandman O;
XX	
XX	WPI: 2001-425874/45.
DR	N-PSDB: AAD09944.
XX	
PT	Drug metabolizing enzymes and encoding polynucleotides, useful for diagnosing, treating and/or preventing autoimmune, inflammatory, cell proliferative, developmental, endocrine, eye, metabolic, and gastrointestinal disorders.
PT	
XX	Claim 1; Page 143-144; 133pp; English.
PS	
XX	
CC	The present sequence is human drug metabolising enzyme (DME-9) protein. Human DME and its nucleic acid molecule are useful for the diagnosis, treatment and prevention of disorders associated with increased or decreased expression of DME. Examples of such disorders include, autoimmune/inflammatory disorder such as acquired immune deficiency syndrome (AIDS), rheumatoid arthritis, osteoporosis; cell proliferative disorder such as actinic keratosis, atherosclerosis; developmental disorder such as epilepsy, anaemia; endocrine disorder such as acromegaly, cretinism, thyrotoxicosis; pancreatic disorder such as diabetes mellitus; eye disorder such as conjunctivitis, glaucoma, iritis; metabolic disorder such as Addison's disease, obesity; gastrointestinal disorder such as anorexia, dysphagia and hepatic tumours including nodular hyperplasia, adenomas and carcinomas. DME DNA is useful for creating 'knockin' humanised animals (pigs) or transgenic animals (mice or rats) to model human disease. DME DNA is also useful in gene therapy. DME and its immunogenic fragments are useful for screening libraries of compounds in several drug screening assays
CC	
XX	Sequence 304 AA;
SQ	

Alignment Scores:

Score:	397.50	Matches:	85
Percent Similarity:	49.3%	Conservative:	53
Best Local Similarity:	30.4%	Mismatches:	91
Query Match:	9.0%	Indels:	51
DB:	7	Gaps:	3
US-10-768-158-1 (1-2419) x ADD12583 (1-261)			
Qy	72	AAGTACTTCGAGTTCCATGGCGTGGCGTCCGCCCTTCTCCGGGGAAGATGAGGAG	131
Db	12	LysLeuLysGluValGluGlyThrLeuLeuGlnProAlaThrValAspAsnTrpSerGln	31
Qy	132	ATCCCAACTCCCGGTGGCGGCCAGGACGTGGATCGTCACCTACCCCAAGTCGGGC	191
Db	32	IleGlnSerPheGluAlaLysProAspAspLeuLeuCysThrTrpProLysAlaGly	51
Qy	192	ACCAGCTTGTCGAGGAGGTGCTACTTGTGTAGCCAGCGGCGTGACCCCGATGAGATC	251
Db	52	ThrThrTrpIleGlnGluIleVal-----	59
Qy	252	GGCTTGATGAACATCGACGACGAGCTCCCGTCTCGGATGACCCACAGCGGGCGCTGGAC	311
Db	60	-----Glu	60
Qy	312	ATCATCAAGAACTGACCTCTCCCGGCTCATCAAGAGCCACCTGCCCTACCGCTTCTGTG	371
Db	61	LysAlaLysAlaMetProSerProArgIleLeuLysThrHisLeuSerThrGlnLeuLeu	80
Qy	372	CCCTCTGACCTCCACAATGAGAGCTCCAAGGTCTATATGGTTCGCAACCCCAAGAT	431
Db	81	ProProSerPheTrpGluAsnAsnCysLysPheLeuTyrValAlaArgAsnAlaLysAsp	100
Qy	432	CTGTGTGTGCTTATTATCACTCCACCGCTCTCTGCGGACCATGACCTACCGAGGCACC	491
Db	101	CysMetValSerTyrTyrHisPheGlnArgMetAsnHisMetLeuProAspProGlyThr	120
Qy	492	TTTCAAGAAATTCCTCCGGAGGTTTATGAATGATTAAGCTGGGCTACGGCTCTGTTGAG	551
Db	121	TrpGluGluTyrPheGluThrPheIleAsnGlyLysValValTrpGlySerTrpPheAsp	140
Qy	552	CACGTGAGGAGTTCTGGGACACCGCATGACATCGAACGTGCTTTTCTCAAGTATGAA	611
Db	141	HisValLysGlyTrpTrpGluMetCysLysAspArgHisGlnIleLeuPheLeuPheTyrGlu	160
Qy	612	GACATGCATCGGACCTGGTGACCATGCTGGAGCAGCTGGCCAGATTCCTGGGGGTGTCC	671
Db	161	AspIleLysArgAspProLysHisGluIleArgLysValMetGlnPheMetGlyLysLys	180
Qy	672	TGTGACAAAGCCCACTGGGAAGCCCTGACGGAGCAC-----	707
Db	181	ValAspGluThrValLeuAspLysIleValGlnGlnGluThrSerPheGluLysMetLysGlu	200
Qy	708	-----TGCCACACGCTGGTGACCGAGTGTGCAACGCT	740
Db	201	AsnProMetThrAsnArgSerThrValSerLysSerIleLeuAspGlnSerIleSerSer	220
Qy	741	GAGGCCCTGCGCGTGGCGGGGAAGAGTTGGGCTGTGGAGGACATCTTCACCGTCTCC	800
Db	221	Phe-----MetArgLysGlyThrValGlyAspTrpLysAsnHisPheThrValAla	237
Qy	801	ATGAATGAGAAGTTGACTGGTGTATAAACAAGATGGGAAAGTGTGACCTCACGTTT	860
Db	238	GlnAsnGluArgPheAspGluIleTyrArgArgLysMetGluGlyThrSerIleAsnPhe	257
RESULT 76			
ADI16670			
ID	ADI16670	standard; protein; 295 AA.	
XX	AC	ADI16670;	
XX	AC		
DT	15-APR-2004	(first entry)	
XX	DE	Human NOVX protein to treat human pathological conditions SeqID206.	

XX	human; NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes; inflammation; autoimmune disorder; allergy; blood disorder; acquired immunodeficiency syndrome; AIDS; obesity; asthma; immunoglobulin (Ig)A nephropathy; cirrhosis; arthritis; Alzheimer's disease; infection; stroke; muscular dystrophy; epilepsy; cytostatic; candidant; antiinflammatory; immunosuppressive; antiallergic; haemostatic; anti-HIV; antidiabetic; antiarteriosclerotic; anorectic; antiasthmatic; nephrotropic; antiarthritic; hepatotropic; neuroprotective; nootropic; antibacterial; virucide; antiparasitic; relaxant; anticonvulsant; neurogenesis; wound healing; angiogenesis; chromosome mapping; tissue typing; pharmacogenomic.
OS	Homo sapiens.
OS	WO200268649-A2.
PN	06-SEP-2002.
PD	
XX	31-JAN-2002; 2002WO-US002785.
PF	
XX	31-JAN-2001; 2001US-0265395P.
PR	31-JAN-2001; 2001US-0265412P.
PR	31-JAN-2001; 2001US-0265514P.
PR	31-JAN-2001; 2001US-0265517P.
PR	02-FEB-2001; 2001US-0266406P.
PR	05-FEB-2001; 2001US-0266767P.
PR	07-FEB-2001; 2001US-0266975P.
PR	07-FEB-2001; 2001US-0267057P.
PR	08-FEB-2001; 2001US-0267459P.
PR	09-FEB-2001; 2001US-0267823P.
PR	15-FEB-2001; 2001US-0268974P.
PR	26-FEB-2001; 2001US-0271664P.
PR	27-FEB-2001; 2001US-0271839P.
PR	27-FEB-2001; 2001US-0271855P.
PR	02-MAR-2001; 2001US-0272788P.
PR	02-MAR-2001; 2001US-0273046P.
PR	14-MAR-2001; 2001US-0275925P.
PR	14-MAR-2001; 2001US-0275947P.
PR	14-MAR-2001; 2001US-0275950P.
PR	14-MAR-2001; 2001US-0275989P.
PR	15-MAR-2001; 2001US-0276448P.
PR	15-MAR-2001; 2001US-0276450P.
PR	16-MAR-2001; 2001US-0276397P.
PR	16-MAR-2001; 2001US-0276768P.
PR	20-MAR-2001; 2001US-0278652P.
PR	26-MAR-2001; 2001US-0278775P.
PR	26-MAR-2001; 2001US-0278778P.
PR	29-MAR-2001; 2001US-0279882P.
PR	29-MAR-2001; 2001US-0279884P.
PR	30-MAR-2001; 2001US-0280147P.
PR	11-APR-2001; 2001US-0282992P.
PR	11-APR-2001; 2001US-0283083P.
PR	20-APR-2001; 2001US-0285133P.
PR	23-APR-2001; 2001US-0285749P.
PR	03-MAY-2001; 2001US-0288327P.
PR	03-MAY-2001; 2001US-0288504P.
PR	29-MAY-2001; 2001US-0294047P.
PR	30-MAY-2001; 2001US-0294473P.
PR	08-JUN-2001; 2001US-0296964P.
PR	18-JUN-2001; 2001US-0298959P.
PR	19-JUN-2001; 2001US-0299324P.
PR	13-AUG-2001; 2001US-0312020P.
PR	16-AUG-2001; 2001US-0312889P.
PR	16-AUG-2001; 2001US-0312908P.
PR	21-AUG-2001; 2001US-0313390P.
PR	28-AUG-2001; 2001US-0315470P.
PR	31-AUG-2001; 2001US-0316447P.
PR	07-SEP-2001; 2001US-0318115P.
PR	07-SEP-2001; 2001US-0318118P.
PR	12-SEP-2001; 2001US-0318740P.
PR	19-SEP-2001; 2001US-0323379P.
PR	18-OCT-2001; 2001US-0330245P.


```

PR 18-OCT-2001; 2001US-0330308P.
PR 14-NOV-2001; 2001US-0332701P.
XX
XX
XX (CURA-) CURAGEN CORP.
XX
XX Tchernev VT, Spytek KA, Zerhusen BD, Patturajan M, Shinkets RA;
XX Li L, Gangolli EA, Radigaru M, Anderson DM, Rastelli L, Miller CE;
XX Gerlach VL, Taupier RJ, Gusev YV, Colman SD, Wolenc AR, Pena CE;
XX Furtak K, Grosse WM, Alsobrook JP, Lepley DK, Rieger DK, Burgess CE;
XX
XX WPI; 2002-706998/76.
XX N-PSDB; ADI16669.
XX
XX New NOVX polypeptides and nucleic acids, useful for preventing or
XX treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
XX atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
XX pharmacogenomics.
XX
XX Claim 1; SEQ ID NO 206; 1498pp; English.
XX
XX This invention relates to a novel nucleic acids, and encoded polypeptides
XX thereof, which have properties related to the stimulation of biochemical
XX or physiological responses in a cell, tissue, organ or organism.
XX Specifically, it refers to the use of biologically active fragments for
XX diagnostic and prognostic assays and furthermore in the treatment of
XX diverse pathological conditions. The present invention describes novel
XX human and murine NOVX proteins, as well as methods to modulate their
XX expression using antisense oligos, ribozymes and peptide nucleic acids.
XX The NOVX polypeptides, polynucleotides and antibodies are useful in
XX treating or preventing NOVX-associated disorders, e.g. cardiomyopathy,
XX atherosclerosis, cancer and diabetes. Furthermore, they may be used in
XX treating or preventing diseases such as inflammation, autoimmune
XX disorders, allergies, blood disorders, acquired immunodeficiency syndrome
XX (AIDS), obesity, asthma, immunoglobulin (IgA) nephropathy, cirrhosis,
XX arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy
XX and epilepsy. Accordingly, these molecules have many activities including
XX cytosstatic, antiand, antiinflammatory, immunosuppressive, antiallergic,
XX haemostatic, anti-HIV, antidiabetic, antiarteriosclerotic, anorectic,
XX antiasthmatic, nephrotropic, antithyroidic, hepatotropic,
XX neuroprotective, nootropic, antibacterial, virucide, antiparasitic,
XX relaxant and anticonvulsant. In addition, they are useful in screening
XX assays to identify small molecules that modulate or inhibit, for example,
XX neurogenesis, wound healing and angiogenesis. The nucleic acids are also
XX used as in chromosome mapping, tissue typing, preventive medicine and
XX pharmacogenomics. This polypeptide is a human NOVX protein of the
XX invention.
XX
XX Sequence 295 AA;
XX
XX Alignment Scores:
XX Pred. No.: 3.09e-28 Length: 295
XX Score: 397.50 Matches: 85
XX Percent Similarity: 54.2% Conservative: 58
XX Best Local Similarity: 32.2% Mismatches: 102
XX Query Match: 9.0% Indels: 19
XX DB: 5 Gaps: 4
XX
XX US-10-768-158-1 (1-2419) x ADI16670 (1-295)
XX
XX QY 126 GAGGAGATCGCAACTTCCCGGTGCGCCGACGACGAGTGTGATCGTCACCTACCCCAAG 185
XX
XX DB 28 GluLysValCysAsnProAlaLysProAspLeuLeuAlaThrTyrProLys 47
XX
XX QY 186 TCCGGCACCACTGTCTCAGAGAGTGTCTACTTGTGTGAGCCAGGGCGCTGACCCCGAT 245
XX
XX DB 48 SerGlyThrThrTyrMetHisGluLeuLeuAspMetIleLeuAsnAspGlyValGlu 67
XX
XX QY 246 GAGATCGCTTGATGAACATCGACGAGCTCCCGGTCTCTGGAG-----TACCCA--- 296
XX
XX DB 68 LysCysLysArgAlaGlnThrLeuAspArgHisAlaPheLeuGluLeuLysPheProHis 87
XX
XX QY 297 -----CAGCGGGGCTCGGACATCATCAAGGAACCTGACCTCTCCCGGCTCATCAAGAGC 350
XX
XX
XX 88 LysGluLysProAspLeuGluPheValLeuGluMetSerSerProGlnLeuLeuLysThr 107
XX
XX QY 351 CACTGCGCCCTACCGCTTTCTGCGCTCTGACCTCACAATGAGAGACTCAAGGTCACTAT 410
XX
XX DB 108 HisLeuProSerHisLeuIleProProSerIleTyrPlysGluAsnCysLysIleValTyr 127
XX
XX QY 411 ATGGCTCGCAACCCCAAGGATCTGGTGTGTCTTATTATCAGTTTCCACCGCTCTCTCGG 470
XX
XX DB 128 ValAlaArgAsnProLysAspCysLeuValSerTyrTyrHisPheHisArgMetAlaSer 147
XX
XX QY 471 ACCATGAGTACCGGACGACCTTTCAAGAATTCTGCGCGAGGTTTATGAATGAATGCTG 530
XX
XX DB 148 PheMetProAspProGlnAsnLeuGluPheTyrGluLysPheMetSerGlyLysVal 167
XX
XX QY 531 GGCTACCGCTCTCGTGTGAGCAGCTCAGAGTTCTGGAGACCCGATGGAGCTCGAAC 590
XX
XX DB 168 ValGlyGlySerTyrPheAspHisMetLysGlyTyrTrpAlaAlaLysAspMetHisArg 187
XX
XX QY 591 GTGCTTTTCTCAAGTATGAAGACATCATCGGACCTGCTGACGATGGTGGACGACGCTG 650
XX
XX DB 188 IleLeuTyrLeuPheTyrGluAspIleLysLysAsnProLysHisGluIleHisLysVal 207
XX
XX QY 651 GCCAGATTCCTGGGGGTGTCTCTGTGACAAGGCCGAGCTGGAAGCCCTGACGAGCAGCTGC 710
XX
XX DB 208 LeuGluPheLeuGluLysThrTyrSerGlyAspValIleAsnLysIleValHisThr 227
XX
XX QY 711 CACCAGCTGGTG-----GACCAGTGTGCAACGCTGAGGCCCTGCC----- 752
XX
XX DB 228 SerPheAspValMetLysAspAsnProMetAlaAsnHisThrAlaValProAlaHisIle 247
XX
XX QY 753 -----GTGGCGCGGGAAGAGTTGGGCTGTGGAAGGACATC 788
XX
XX DB 248 PheAsnHisSerIleSerLysPheMetArgLysGlyMetProGlyAspTyrLysAsnHis 267
XX
XX QY 789 TTCACCTCTCCATGAATGAGAGTTTGACTTGTGTATAAACAAGAAGATGGGAAAGTGT 848
XX
XX DB 268 PheThrValAlaMetAsnGluAsnPheAspLysHisTyrGluLysLysMetAlaGlySer 287
XX
XX QY 849 GACTCTACGTTT 860
XX
XX DB 288 ThrLeuAsnPhe 291
XX
XX RESULT 77
XX ADN42324
XX ID ADN42324 standard; protein; 295 AA.
XX
XX AC ADN42324;
XX
XX XX 17-JUN-2004 (first entry)
XX
XX DE Human novel proteinNOV 56b.
XX
XX KW Human; NOVX; cancer; diabetes; cardiomyopathy; atherosclerosis.
XX
XX OS Homo sapiens.
XX
XX PN US2004033493-A1.
XX
XX PD 19-FEB-2004.
XX
XX PF 31-JAN-2002; 2002US-00072012.
XX
XX PR 31-JAN-2001; 2001US-0265395P.
XX
XX PR 31-JAN-2001; 2001US-0265412P.
XX
XX PR 31-JAN-2001; 2001US-0265514P.
XX
XX PR 02-FEB-2001; 2001US-0265517P.
XX
XX PR 05-FEB-2001; 2001US-0266406P.
XX
XX PR 07-FEB-2001; 2001US-0266767P.
XX
XX PR 07-FEB-2001; 2001US-0266975P.
XX
XX PR 08-FEB-2001; 2001US-0267057P.
XX
XX PR 09-FEB-2001; 2001US-0267459P.
XX
XX PR 15-FEB-2001; 2001US-0268974P.

```



```

Db 141 ProllysAspCysLeuValSerTyrTrpHisPheHisArgMetAlaSerPheMetProAsp 160
Qy 483 CGAGGCACCTTCAAGAAATTCGCGGAGGTTTATGAATGATAAGCTGGCTACGGCTCC 542
Db 161 ProGlnAsnLeuGluGluPheTyrGluLysPheMetSerGlyLysGlyGluPheGlySer 180
Qy 543 TGGTTTGAGCAGTCGACGAGATTCCTGGGAGCACCGCATGGACTCGAACTGCTTTTCTC 602
Db 181 TrpPheAspHisValLysGlyTrpTrpAlaAlaLysAspMetHisArgIleLeuTyrLeu 200
Qy 603 AAGTATGAGACATGCATCGGACCTGCTGACGATGCTGGAGCAGCTGGCCAGATTCCTG 662
Db 201 PheTyrGluAspIleLysGlnAsnProLysHisGluIleHisLysValLeuGluPheLeu 220
Qy 663 GGGGTGCTCTGACAGCGCCAGCTGGAAGCCCTGACGGAGCACCTGCCACCATGCTGTG 722
Db 221 GluLysThrTrpSerGlyAspValIleAsnLysIleValHisHisThrSerPheAspVal 240
Qy 723 -----GACCATGTGTCAGCTGAGCCCTGCGCCGTG----- 755
Db 241 MetLysAspAsnProMetAlaAsnHisThrAlaValProAlaHisIlePheAsnHisSer 260
Qy 756 -----GGCCGGGAGAGTTGGCTGTGGAAGGACATCTTCACCGTC 797
Db 261 IleSerLysPheMetArgLysGlyGlyMetProGlyAspTrpLysAsnHisPheThrVal 280
Qy 798 TCCATGATGAGAAGTTTGACTTGTGTATAAACAGAGATGGGAAAGTGTGACCTCACG 857
Db 281 AlaLeuAsnGluAsnPheAspLysHisTyrGluLysLysMetAlaGlySerThrLeuAsn 300
Qy 858 TTT 860
Db 301 Phe 301

RESULT 79
AAE38814
ID AAE38814 standard; protein; 305 AA.
XX
AC AAE38814;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human sulfotransferase-like protein (POLY13).
XX
KW Human; tumour; inflammatory disorder; vaccine; gene therapy; cytostatic;
KW sulfotransferase-like protein; POLY13; chromosome 2.
XX
OS Homo sapiens.
XX
PN US2003050232-A1.
XX
PD 13-MAR-2003.
XX
PF 19-APR-2001; 2001US-00839446.
XX
PR 19-APR-2000; 2000US-0198293P.
PR 20-APR-2000; 2000US-0198645P.
PR 25-APR-2000; 2000US-0198476P.
PR 26-APR-2000; 2000US-0199880P.
PR 26-APR-2000; 2000US-0200024P.
PR 09-JUN-2000; 2000US-0200025P.
PR 17-JUL-2000; 2000US-0210809P.
PR 11-AUG-2000; 2000US-0224610P.
PR 27-FEB-2001; 2001US-0271814P.
XX
(PADU) TAUPIER R J.
(PADI) PADIGARU M.
(PVT) SPYTEK K A.
(BURG) BURGESS C E.
(VERN) VERNET C A M.
(FERN) FERNANDES E R.

```

```

PA (SHIM/) SHIMKETS R A.
PA (LIUX/) LIU X.
PA (MAJU/) MAJUMDER K.
PA (COLM/) COLMAN S D.
PA (ZERH/) ZERHUSEN B D.
XX
PI Taupier RJ, Padigaru M, Spytek KA, Burgess CE, Vernet CM;
PI Fernandes ER, Shimkets RA, Liu X, Majumder K, Colman SD;
PI Zerhusen BD;
XX
DR WPI: 2003-605764/57.
DR N-PSDB: AAD58965.
XX
PT New POLYX nucleic acid, useful for preparing a composition for treating
PT or preventing e.g., tumor or inflammatory disorder.
XX
PS Claim 1; Page 35; 75pp; English.
XX
CC The invention relates to new POLYX nucleic acid useful for preparing a
CC composition for treating or preventing tumour or inflammatory disorder.
CC The invention is useful as vaccine and in gene therapy. The nucleic acid
CC is useful for preparing a composition for treating or preventing e.g.,
CC tumour or inflammatory disorder. The present sequence is human
CC sulfotransferase-like protein (POLY13). POLY13 gene is located on
CC chromosome 2. Note: The present sequence is encoded by a DNA containing
CC translational exceptions
XX
SQ Sequence 305 AA;
XX
Alignment Scores:
Pred. No.: 3.91e-28 Length: 305
Score: 396.50 Matches: 93
Percent Similarity: 50.8% Conservative: 60
Best Local Similarity: 30.9% Mismatches: 127
Query Match: 9.0% Indels: 21
DB: 7 Gaps: 5
XX
US-10-768-158-1 (1-2419) x AAE38814 (1-305)
Qy 21 ATGGCGGAGCGAGCGCGGAGACCCCGAGACCCCGGGGAGTTCGAG---AGCAAGTAC 77
Db 1 MetAlaLysIleGluLysAsnAlaProThrMetGluLysLysProGluLeuPheAsnIle 20
Qy 78 TTCGAGTTCATGCGGTGCGGCTGCCCGCTTTCGCGCGGAGAGATGGAGGAGATGCC 137
Db 21 MetGluValAspGlyValProThrLeuIleLeuSerLysGluTrpTrpGluLysValCys 40
Qy 138 AACTTCCCGTGGCGGCGGAGCGGTGGATCGTCACCTACCCAGTCCGGCACCAGC 197
Db 41 AsnPheGlnAlaLysProAspAspLeuIleLeuAlaThrTyrProLysSerGlyThr 60
Qy 198 TTGCTGCAGGAGGTGGTCTACTTGGTGACGCGGCGCTGACCCCGATGAGATCGGCTG 257
Db 61 TrpMetHisGluIleLeuAspMetIleLeuAsnAspGlyAspValGluLysCysLysArg 80
Qy 258 ATGAACATCGACGAGCAGCTCCCGGTCTCGAG-----TACCCA-----CAGCCG 302
Db 81 AlaGlnThrLeuAspArgHisAlaPheLeuGluLysPheProHisLysGluLysPro 100
Qy 303 GGCTTGACATCATCAGGAACCTGACCTTCCCGGCTCATCAGAGCCACCTGCCCTAC 362
Db 101 AspLeuGluPheValLeuGluMetSerSerProGlnLeuIleLysThrHisLeuProSer 120
Qy 363 CGCTTTCTGCGCTCTGACCTCCCAATGGAGACTCCAAATGGAGCTCATCTATATGGCTCGCAAC 422
Db 121 HisLeuIleProProSerIleTrpLysGluAsnCysLysIleValTrpValAlaArgAsn 140
Qy 423 CCCAAGGATCTGGTGGTGTCTTATATACCTTCCACCGCTCTCTGCGGACCATGAGTAC 482
Db 141 ProllysAspCysLeuValSerTyrTrpHisPheHisArgMetAlaSerPheMetProAsp 160
Qy 483 CGAGGCACCTTCAAGAAATTCGCGGAGGTTTATGAATGATGATGAGTGGCTACGGCTCC 542

```

Db 161 ProGlnAenLeuGluPheTyrGluLysPheMetSerGlyLysGlyGluPheGlySer 180
 Qy 543 TGGTTTGAGCAGCGTCGAGGAGTTCTGGAGACCGCATGGAAGTCTGAGTCTTTTCTC 602
 Db 181 TrpPheAepHisValLysGlyTrpAlaAlaLysAepMetHisAArgileLeuTyrLeu 200
 Qy 603 AAGTATGAGCATGTCATCGGAGCCTGCTGACGATGGTGGAGCAGCTGGCCAGATTCTCTG 662
 Db 201 PheTyrGluAepIleLysGlnAenProLysHisGluIleHisLysValLeuGluPheLeu 220
 Qy 663 GGGGTGCTCTGTGACAAAGCCAGCTGGAAAGCCCTGACGAGCAGCTGCCACCACTGGTG 722
 Db 221 GluLysTrpSerGlyAepValIleAenLysIleValHisHisTrpSerPheAepVal 240
 Qy 723 -----GACCAGTCTGCAAGCTGAGCGCTGAGCGCTGCCCTGCGCTG----- 755
 Db 241 MetLysAepAenProMetAlaAenHisThrAlaValProAlaHisIlePheAenHisSer 260
 Qy 756 -----GGCCGGGGAAGAGTTGGGCTGTGGAAGGACATCTTCACCGTC 797
 Db 261 IleSerLysPheMetArgLysGlyGlyMetProGlyAepTrpLysAenHisPheThrVal 280
 Qy 798 TCCATGAATGAGAAAGTTTGACTTGGTGTATTAACAGAGATGGGAAAGTGTGACCTCAGC 857
 Db 281 AlaLeuAenGluAenPheAepLysHisTyrGluLysLysMetAlaGlySerThrLeuAen 300
 Qy 858 TTT 860
 Db 301 Phe 301
 RESULT 80
 ADRI19665
 ID ADRI19665 standard; protein; 297 AA.
 XX AC ADRI19665;
 XX AC
 DT 07-OCT-2004 (first entry)
 XX AC
 DE Human drug metabolising enzyme (DME) -4 protein sequence.
 XX AC
 KW drug metabolising enzyme; DME; cytostatic; immunosuppressive;
 KW antiinflammatory; endocrine; ophthalmological; gastrointestinal;
 KW hepatotropic; cancer; cell proliferative disorder; autoimmune disorder;
 KW inflammatory disorder; endocrine disorder; eye disorder;
 KW gastrointestinal disorder; liver disorder; metabolic disorder; enzyme;
 KW human.
 XX OS Homo sapiens.
 XX OS
 XX WO200226988-A2.
 XX PN
 XX PD 04-APR-2002.
 XX PF
 XX PF 28-SEP-2001; 2001WO-US030662.
 XX PR 29-SEP-2000; 2000US-0236947P.
 XX PR 06-OCT-2000; 2000US-0238864P.
 XX PR 20-OCT-2000; 2000US-0242323P.
 XX PR 09-NOV-2000; 2000US-0247581P.
 XX PR 16-NOV-2000; 2000US-0249519P.
 XX PR 22-NOV-2000; 2000US-0252834P.
 XX PR 30-NOV-2000; 2000US-0250567P.
 XX (INCY-) INCYTE GENOMICS INC.
 XX PA
 XX PI Azimzai Y, Baughin MR, Borowsky ML, Ding L, Duggan BW,
 PI Elliott VS, Gandhi AR, Griffin JA, Hafalia AJA, Ison CH, Khan FA;
 PI Lal P, Lee EA, Lu DAM, Nguyen DB, Arvizu C, Policky JL, Ramkumar J;
 PI Ring HZ, Sanjanwala MS, Tang YT, Tribouley CM, Narinder WK;
 PI Walsh RT, Warren BA, Xu Y, Yang J, Yao MG, Yue H;
 XX WPI; 2002-362498/39.
 DR N-PSDB; ADRI19683.
 DR

XX Human drug metabolizing enzymes, useful in the diagnosis and treatment of
 PT disorders associated with aberrant (DME) activity, e.g., cancer and
 PT autoimmune disorders.
 XX
 PS Claim 1; SEQ ID NO 4; 142pp; English.
 XX
 CC This invention relates to novel drug metabolising enzymes (DME) and the
 CC nucleotide sequences which encode them. The invention may be useful for
 CC the development of compounds with a cytostatic, immunosuppressive,
 CC antiinflammatory, endocrine, ophthalmological, gastrointestinal or
 CC hepatotropic activity acting as an agonist or antagonist of drug
 CC metabolising enzyme activity. The invention may be used in the diagnosis
 CC and treatment of disorders associated with decreased or increased
 CC expression or activity of drug metabolising enzymes. Such disorders
 CC include cancer, cell proliferative disorders, autoimmune/inflammatory,
 CC endocrine, eye, gastrointestinal (including liver disorders) and
 CC metabolic disorders. The present sequence is that of a human drug
 CC metabolising enzyme (DME) of the invention. Note: This sequence did not
 CC form part of the printed specification but was obtained in electronic
 CC format from EPO.
 XX
 SQ Sequence 297 AA;
 Alignment Scores:
 Pred. No.: 4.32e-27 Length: 297
 Score: 385.50 Matches: 84
 Percent Similarity: 53.7% Conservative: 55
 Best Local Similarity: 32.4% Mismatches: 105
 Query Match: 8.8% Indels: 15
 DB: Gaps: 3
 US-10-768-158-1 (1-2419) x ADRI19665 (1-297)
 Qy 126 GAGGAGATCGCCAACTTCCCGGTGCGGCCGAGCGAGTGGATCGTACACTACCCCAAG 185
 Db 36 AsplysiletrpAenPheGlnAlaLysProAepAplLeuIleSerThrTyrProLys 55
 Qy 186 TCGGGACACGCTGCTGCGAGGAGGTGCTACTTGTGTAGCCAGGGCGGTGACCCCGAT 245
 Db 56 AlaGlyThrTrpMetHisGluIleLeuAepMetIleLeuAenAepGlyAepValGlu 75
 Qy 246 GAGATCGGCTTGATGAATCGACGACGAGCTCCCGCTCTGGAGTACCCACACCCGGCG 305
 Db 76 LysCysLysArgAlaGlnThrLeuAepArgHisAlaPheLeuGluLysLysPhePro--- 94
 Qy 306 CTGACATCATCAAGAACTGACCTCTCCCGCTCATCAAGAGCCACTGCTCCCTACCGC 365
 Db 95 LeuGluPheValLeuGluMetSerSerProGlnLeuIleLysThrHisLeuProSerHis 114
 Qy 366 TTTCTGCGCTCTGACCTCCACATGGAGACTCCCAAGTCTCTCTCGGACCATGACTACCGA 425
 Db 115 LeuIleProProSerIleTrpLysGluAenCysLysIleValTyrValAlaAArgAenPro 134
 Qy 426 AAGGATCTGTGTGTCTTATTATCAGTTCCACCGCTCTCTCGGACCATGACTACCGA 485
 Db 135 LysAepCysLeuValSerTyrHisPheHisArgMetAlaSerPheMetProAepPro 154
 Qy 486 GGCACCTTCAAGAACTCTGCGGAGGTTTATGATGATAGCTGGGTACGGCTCTCTGG 545
 Db 155 GlnAenLeuGluPheTyrGluLysPheMetSerGlyLysValValGlyArgSerTrp 174
 Qy 546 TTTGAGCAGCTGCGAGGAGTTCTGGAGCACCGCATGGACTCGAAGCTGCTTTTCTCAAG 605
 Db 175 PheAepHisValLysGlyTrpTrpAlaLysAepThrHisArgIleLeuTyrLeuPhe 194
 Qy 606 TATGAGACATGTCATCGGAGCTGCTGGTACGATGTGGAGCAGCTGGCCAGATTCCTGGGG 665
 Db 195 TyrGluAepIleLysLysAenProLysHisGluIleHisLysValLeuGluPheLeuGlu 214
 Qy 666 GTGTCCTGTGACAAAGGCCAGCTGGAAGCCCTGACGAGGACACTGCCACCACTGGTG--- 722
 Db 215 LysThrLeuSerGlyAepValIleAenLysIleValHisHisThrSerPheAepValMet 234

Qy 723 -----GACCAAGTCTGCAACGCTGAGCGCCCTGCCC----- 752
Db 235 LysAspAsnProMetAlaAsnHisThrAlaValProAlaHisIlePheAsnHisSerIle 254
Qy 753 -----GTGGCGCGGGAAGAGTTGGGCTGTGGAGGACATCTTCACCGTCTCCATG 803
Db 255 SerLysPheMetArgLysGlyMetProGlyAspTrpLysAsnHisPheThrValAlaMet 274
Qy 804 AATGAGAAGTTTGACTTGGTGTATATAACAGAAAGATGGAAAGTGTGACCTCACGTTT 860
Db 275 AsnGluAsnPheAspLysHisTyrGluLysLysMetAlaGlySerThrLeuAsnPhe 293
RESULT 81
ABM84008
ID ABM84008 standard; protein; 262 AA.
XX
AC ABM84008;
XX
DT 18-NOV-2004 (first entry)
XX
XX Human diagnostic and therapeutic pprotein SEQ ID NO:4257.
XX gene therapy; human diagnostic and therapeutic polynucleotide; dithp.
XX Homo sapiens.
XX
XX WO2004023973-A2.
XX
XX 25-MAR-2004.
XX
XX 12-SEP-2003; 2003WO-US028227.
XX
XX 12-SEP-2002; 2002US-0410259P.
XX
XX 12-SEP-2002; 2002US-0410260P.
XX
XX (INCY-) INCYTE CORP.
XX
XX Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
XX Harthorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
XX Mooney EM, Deleane AM, Panesar IS, Banville SC, Reddy TP;
XX Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstein EH;
XX Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LW;
XX Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;
XX Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
XX Patury S, Shi X, Suarez CJ;
XX
XX WPI; 2004-329368/30.
XX
XX N-PSDB; ACN42660.
XX
XX New diagnostic and therapeutic polynucleotides and polypeptides, useful
XX in diagnosing a condition, disease or disorder associated with human
XX molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
XX in gene mapping.
XX
XX Claim 27; Page; 190pp; English.
XX
XX The invention relates to novel diagnostic and therapeutic polynucleotides
XX selected from one of the 2722 sequences defined in the specification. A
XX polynucleotide of the invention may have a use in gene therapy. The human
XX diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
XX used to diagnose a particular condition, disease or disorder associated
XX with human molecules, e.g. cell proliferative disorders,
XX autoimmune/inflammatory disorder, developmental disorder, endocrine
XX disorder, neurological disorders, gastrointestinal disorders, or
XX infections caused by virus, bacteria, fungi or parasite. The dithp
XX molecules may also be used in genetic mapping, in identifying individuals
XX from minute biological samples, in detecting single nucleotide
XX polymorphisms, as molecular weight markers, and for somatic or germline
XX gene therapy. The present sequence represents a dithp protein of the
XX invention. Note: The sequence data for this patent is not represented in
XX the printed specification, but was obtained in electronic format directly
XX from WIPO at www.wipo.int/pct/en/sequences/listing.htm

XX SQ Sequence 262 AA;
Alignment Scores:
Pred. No.: 5,04e-27 Length: 262
Score: 384,50 Matches: 84
Percent Similarity: 48.4% Conservative: 49
Best Local Similarity: 30.5% Mismatches: 91
Query Match: 8.7% Indels: 51
DB: Gaps: 4
US-10-768-158-1 (1-2419) x ABM84008 (1-262)
Qy 90 GCGCTGCGCTGCGCGCCCTTCTGCGCGGGAAGATGAGGAGATGCGCAACTTCCCGGTG 149
Db 17 GlyValProLeuIleLysTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla 36
Qy 150 CGGCCAGCGACGCTGTCGATCGTCACCTACCCCAAGTCCGCGCACCGCTTCTCGCAGG 209
Db 37 ArgProAspAspLeuLeuIleAsnThrTyrProLysSerGlyThrTrpValSerGln 56
Qy 210 GTGCTCTACTTGTGAGCCAGCGCTGACCCCGATGAGATCGGCTTGCATGAACATCGAC 269
Db 57 IleLeuAspMetIleTyrGlnGlyAspLeuGluLysCysAsnArgAlaProIleTyr 76
Qy 270 GAGCAGTCCCGTCTCGAGTACCCACAGCCGCGCTCGACATCATCAAGAACTGACC 329
Db 77 ValArgValProPheLeuGluValAsnAspPro----- 87
Qy 330 TCTCCCGCTCATCAAGAGCCACCTGCCCTTCTGCGCTTCTGACCTCCACNAT 389
Db 87 ----- 87
Qy 390 GGAGACTCCAAGGTCATCTATATGCTCGCAACCCCAAGGATCTGGTGGTCTTATTAT 449
Db 88 GlyGluProSerGlyValTyrValAlaArgAsnProLysAspValAlaValSerTyr 107
Qy 450 CAGTTCCACCGCTCTCTGCGGACCATGAGTACCGAGGACCTTTTCAAGAAATCTCGCGG 509
Db 108 HisPheHisArgMetGluLysAlaHisProGluProGlyThrTrpAspSerPheLeuGlu 127
Qy 510 AGGTTTATGAATGATAGCTGGGCTACGGTCTCGCTTTCAGCAGCTGCGAGGCTCTGG 569
Db 128 LysPheMetAlaGlyGluValSerTyrGlySerTrpTyrGlnHisValGlnGluTrpTrp 147
Qy 570 GAGCACCGCATGGACTCGAAGCTCTTTCTCAAGTATGAAGACATGCATCGGAGACCTG 629
Db 148 GluLeuSerArgThrHisProValLeuTyrLeuPheTyrGluAspMetLysGluAsnPro 167
Qy 630 GTGACGATGTGGAGCAGCTGGCCAGATTCTGGGGGTGCTCTGTGACAAGGCCAGCTG 689
Db 168 LysArgGluIleGlnLysIleLeuGluPheValGlyArgSerLeuProGluGluThrMet 187
Qy 690 GAAGCCCTGAGGAGCACTGC----- 710
Db 188 AspPheMetValGlnHisThrSerPheLysGluMetLysLysAsnProMetThrAsnTyr 207
Qy 711 -----CACACGCTGGTGACAGCTGCTGCAACGCTGAGGCGCTGCC-----GTG 755
Db 208 ThrThrValProGlnGluLeuMetAspHisSerIleSer-----ProPheMet 223
Qy 756 GGCCGGGAAGAGTTGGGCTGTGGAAGGACATCTTCACCGCTCTCCATGAAGAAGTTT 815
Db 224 ArgLysGlyMetAlaGlyAspTrpLysThrPheThrValAlaGlnAsnGluArgPhe 243
Qy 816 GACTTGTGTATAAACAGAGATGGGAAAGTGTGACCTCAGCTT 860
Db 244 AspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe 258
RESULT 82
ABM82616
ID ABM82616 standard; protein; 262 AA.
XX

```

AC ABM82616;
XX
XX 18-NOV-2004 (first entry)
XX
XX Human diagnostic and therapeutic pprotein SEQ ID NO:2865.
XX
XX gene therapy; human diagnostic and therapeutic polynucleotide; dithp.
XX
XX Homo sapiens.
XX
XX WO2004023973-A2.
XX
XX 25-MAR-2004.
XX
XX 12-SEP-2003; 2003WO-US028227.
XX
XX 12-SEP-2002; 2002US-0410259P.
XX
XX 12-SEP-2002; 2002US-0410260P.
XX
XX (INCY-) INCYTE CORP.
XX
XX Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
XX Harthorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
XX Mooney EM, Deleane AM, Panesar IS, Banville SC, Reddy TP;
XX Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
XX Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LI;
XX Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirkon ES;
XX Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
XX Patury S, Shi X, Suarez CJ;
XX
XX WPI: 2004-329368/30.
XX
XX N-PSDB; ACN41268.
XX
XX New diagnostic and therapeutic polynucleotides and polypeptides, useful
XX in diagnosing a condition, disease or disorder associated with human
XX molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
XX in gene mapping.
XX
XX Claim 27; Page: 190pp; English.
XX
XX The invention relates to novel diagnostic and therapeutic polynucleotides
XX selected from one of the 7222 sequences defined in the specification. A
XX polynucleotide of the invention may have a use in gene therapy. The human
XX polynucleotide and therapeutic polynucleotides (dithp) or polypeptides may be
XX used to diagnose a particular condition, disease or disorder associated
XX with human molecules, e.g. cell proliferative disorders,
XX autoimmune/inflammatory disorder, developmental disorder, endocrine
XX disorder, neurological disorders, gastrointestinal disorders, or
XX infections caused by virus, bacteria, fungi or parasite. The dithp
XX molecules may also be used in genetic mapping, in identifying individuals
XX from minute biological samples, in detecting single nucleotide
XX polymorphisms, as molecular weight markers, and for somatic or germline
XX gene therapy. The present sequence represents a dithp protein of the
XX invention. Note: The sequence data for this patent is not represented in
XX the printed specification, but was obtained in electronic format directly
XX from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX
XX SQ Sequence 262 AA;
XX
XX Alignment Scores:
XX Pred. No.: 5,04e-27 Length: 262
XX Score: 384.50 Matches: 84
XX Percent Similarity: 48.4% Conservative: 49
XX Best Local Similarity: 30.8% Mismatches: 91
XX Query Match: 8.7% Indels: 51
XX DB: 8 Gaps: 4
XX
XX US-10-768-158-1 (1-2419) x ABM82616 (1-262)
XX
XX 90 GCGTGGCGGTCGCCCTTCTGCGGGGAGAGATGGAGGATCGCACTTCCCGTG 149
XX
XX 17 GlyValProLeuIleLysTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla 36

```

```

QY 150 CGGCCAGCAGCGTGTGGATCGTCACCTACCCAAAGTCGCGCACACGCTTGCTCAGGAG 209
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 37 ArgProAspLeuLeuIleAsnThrTyrProLysSerGlyThrThrTrpValSerGln 56
QY 210 GTGGTCTACTTGGTAGCCAGCGCTGACCCCGATGAGATCGGCTTGTATGAACATCGAC 269
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 57 IleLeuAspMetIleTyrGlnGlyGlyAspLeuGluLysCysAsnArgAlaProIleTyr 76
QY 270 GAGCAGCTCCGGTCTCGAGTACCCACAGCGCGGCTCGACATCATCAAGGAAGTACC 329
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 77 ValArgValProPheLeuGluValAsnAspPro----- 87
QY 330 TCTCCCGCGCTCATCAAGAGCCACCTGCGCTTCTGCGCTTCTGCGCTCTGACCTCCACAAT 389
DB ----- 87
XX 390 GGAGACTCCAAGTCTATATATGGTTCGCAACCCCAAGGATCTGGTGTCTTATTAT 449
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 88 GlyGluProSerGlyValTyrValAlaArgAsnProLysAspValAlaValSerTyr 107
QY 450 CAGTTCACCGCTCTCGCGGACCATGAGTACCGAGGACCTTTCAGAAATTCGCGG 509
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 108 HisPheHisArgMetGluLysAlaHisProGluProGlyThrTrpAspSerPheLeuGlu 127
QY 510 AGTTTATGAATGATAGTGGCTACGGCTCTCTGGTTTTCAGCACGTGCAGGAGTCTCGG 569
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 128 LysPheMetAlaGlyGluValSerTyrGlySerTyrGlnHisValGlnGluTrpIrp 147
QY 570 GAGCAGCCGCTGAGTACGAACTGCTTTTCTCAAGTATGAAGACATGCGGACCTG 629
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 148 GluLeuSerArgThrHisProValLeuTyrLeuPheTyrGluAspMetLysGluAsnPro 167
QY 630 GTACAGTGTGTGAGCAGTGGCCAGATTCTCTGGGTGTCTGTGACAAAGGCCACCTG 689
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 168 LysArgGluIleGlnLysIleLeuGluPheValGlyArgSerLeuProGluGluThrMet 187
QY 690 GAAGCCCTGACGAGCAGTGC----- 710
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 188 AspPheMetValGlnHisThrSerPheLysGluMetLysLysAsnProMetThrAsnTyr 207
QY 711 -----CACCACTGGTGGACAGTGTGCAACCGCTGAGGCCCTGCCC---GTG 755
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 208 ThrThrValProGlnGluLeuMetAspHisSerIleSer-----ProPheMet 223
QY 756 GCGCGGGAAGAGTGGCTGTGGAAGACATCTTCAACCGCTTCCATGATGAAGTCTT 815
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 224 ArgLysGlyMetAlaGlyAspTyrLysThrPheThrValAlaGlnAsnGluArgPhe 243
QY 816 GACTTGGTGTATAAACAAGAGATGGGAAAGTGCACCTCAGCTT 860
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 244 AspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe 258
XX
XX RESULT 83
XX ABM84005
XX ID ABM84005 standard; protein; 262 AA.
XX
XX AC ABM84005;
XX
XX 18-NOV-2004 (first entry)
XX
XX Human diagnostic and therapeutic pprotein SEQ ID NO:4254.
XX
XX gene therapy; human diagnostic and therapeutic polynucleotide; dithp.
XX
XX Homo sapiens.
XX
XX WO2004023973-A2.
XX
XX 25-MAR-2004.
XX
XX 12-SEP-2003; 2003WO-US028227.
XX
XX 12-SEP-2002; 2002US-0410259P.
XX
XX 12-SEP-2002; 2002US-0410260P.
XX
XX (INCY-) INCYTE CORP.
XX
XX Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
XX Harthorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
XX Mooney EM, Deleane AM, Panesar IS, Banville SC, Reddy TP;
XX Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
XX Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LI;
XX Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirkon ES;
XX Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
XX Patury S, Shi X, Suarez CJ;
XX
XX WPI: 2004-329368/30.
XX
XX N-PSDB; ACN41268.
XX
XX New diagnostic and therapeutic polynucleotides and polypeptides, useful
XX in diagnosing a condition, disease or disorder associated with human
XX molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
XX in gene mapping.
XX
XX Claim 27; Page: 190pp; English.
XX
XX The invention relates to novel diagnostic and therapeutic polynucleotides
XX selected from one of the 7222 sequences defined in the specification. A
XX polynucleotide of the invention may have a use in gene therapy. The human
XX polynucleotide and therapeutic polynucleotides (dithp) or polypeptides may be
XX used to diagnose a particular condition, disease or disorder associated
XX with human molecules, e.g. cell proliferative disorders,
XX autoimmune/inflammatory disorder, developmental disorder, endocrine
XX disorder, neurological disorders, gastrointestinal disorders, or
XX infections caused by virus, bacteria, fungi or parasite. The dithp
XX molecules may also be used in genetic mapping, in identifying individuals
XX from minute biological samples, in detecting single nucleotide
XX polymorphisms, as molecular weight markers, and for somatic or germline
XX gene therapy. The present sequence represents a dithp protein of the
XX invention. Note: The sequence data for this patent is not represented in
XX the printed specification, but was obtained in electronic format directly
XX from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX
XX SQ Sequence 262 AA;
XX
XX Alignment Scores:
XX Pred. No.: 5,04e-27 Length: 262
XX Score: 384.50 Matches: 84
XX Percent Similarity: 48.4% Conservative: 49
XX Best Local Similarity: 30.8% Mismatches: 91
XX Query Match: 8.7% Indels: 51
XX DB: 8 Gaps: 4
XX
XX US-10-768-158-1 (1-2419) x ABM82616 (1-262)
XX
XX 90 GCGTGGCGGTCGCCCTTCTGCGGGGAGAGATGGAGGATCGCACTTCCCGTG 149
XX
XX 17 GlyValProLeuIleLysTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla 36

```


PT in diagnosing a condition, disease or disorder associated with human
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
XX in gene mapping.

PS Claim 27; Page: 190pp; English.

XX
XX
CC The invention relates to novel diagnostic and therapeutic polynucleotides
CC selected from one of the 2722 sequences defined in the specification. A
CC polynucleotide of the invention may have a use in gene therapy. The human
CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
CC used to diagnose a particular condition, disease or disorder associated
CC with human molecules, e.g. cell proliferative disorders,
CC autoimmune/inflammatory disorder, developmental disorders, endocrine
CC disorder, neurological disorders, gastrointestinal disorders, or
CC infections caused by virus, bacteria, fungi or parasite. The dithp
CC molecules may also be used in genetic mapping, in identifying individuals
CC from minute biological samples, in detecting single nucleotide
CC polymorphisms, as molecular weight markers, and for somatic or germline
CC gene therapy. The present sequence represents a dithp protein of the
CC invention. Note: The sequence data for this patent is not represented in
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX

SQ Sequence 253 AA;

Alignment Scores:

Pred. No.:	6.87e-27	Length:	253
Score:	383.00	Matches:	83
Percent Similarity:	47.2%	Conservative:	45
Best Local Similarity:	30.6%	Mismatches:	91
Query Match:	8.7%	Indels:	52
DB:	8	Gaps:	3

US-10-768-158-1 (1-2419) x ABM82696 (1-253)

Qy	90	GGCGTGGCGTGGCGCGCTTCTGCGCGGGAAGATGGAGGATCGCCAACTTCCGGTG	149
Db	17	GlyValProLeuIleLeuTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla	36
Qy	150	CGGCCACGACGAGTGTGGATCGTCACCTACCCCAAGTCCGGACCGAGTTCTGCGAGAG	209
Db	37	ArgProAspLeuLeuIleSerThrTyrProLysSer-----	49
Qy	210	GTGCTCTACTTGGTGGCAGCGGCGCTGACCCCGATGAGATCGGCTTGATGAACATCGAC	269
Db	49	-----	49
Qy	270	GAGCAGCTCCCGTCTGGAGTACCCACAGCGGGCGCTGGACATCATCAAGGAACGTGACC	329
Db	50	-----GlyMetGluThrLeuLysAspThrPro	58
Qy	330	TCTCCCGCTCATCAAGAGCACCTGCCCTTCTGCGCTTCTGACCTCCACAAAT	389
Db	59	AlaProArgLeuLeuLysThrHisLeuProLeuAlaLeuLeuProGlnThrLeuLeuAsp	78
Qy	390	GGAGACTCCAGGTATCATATGCTCGCAACCCCAAGGATCGTGGTGTCTTATTTAT	449
Db	79	GlnLysValLysValValTyrValAlaArgAsnAlaLysAspValAlaValSerTyrTyr	98
Qy	450	CAGTTCCACCGCTCTCTCGGACCATGAGCTACCGGACCTTTCAAGAAATCTGCGCGG	509
Db	99	HisPheTyrHisMetAlaLysValHisProGluProGlyThrTrpAspSerPheLeuGlu	118
Qy	510	AGGTTTATGAATGATAAGCTGGGTACCGCTCTGTTTGGACGACGTGACAGAGTTCTGG	569
Db	119	LysPheMetValGlyGluValSerTyrGlySerTrpTyrGlnHisValGlnGluTrpTrp	138
Qy	570	GAGCACCCATGGACTCCAACTGCTTTTCTCAAGTATGAGACATCATCGGACCTG	629
Db	139	GluLeuSerArgThrHisProValLeuTyrLeuPheTyrGluAspMetLysGluAsnPro	158
Qy	630	GTGACGATGGTGGACGACCTGCCAGATTCTCTGGGGGTGTCCTGTGACAAGGCCACGCTG	689

Db	159	LysArgGluIleGlnLysIleLeuGluPheValGlyArgSerLeuProGluGluThrVal	178
Qy	690	GAAGCCCTGACGAGGAGCTGC-----CACCAGCTGGTGACCACTGCTGCAACGCT	740
Db	179	AspPheValGlnHisThrSerPheLysGluMetLysLysAsnProMetThrAsnTyr	198
Qy	741	GAGCCCTGCGC-----GTGGGCGCGGGAAGA	767
Db	199	ThrThrValProGlnGluPheMetAspHisSerIleSerProPheMetArgLysGlyMet	218
Qy	768	GTGGGCTGTGGAGGACATCTTCCACCGTCTCCATGAATGAGAGATTTGACTTGGTGTAT	827
Db	219	AlaGlyAspTrpLysThrThrPheThrValAlaGlnAsnGluArgPheAspAlaAspTyr	238
Qy	828	AAACAGAGATGGAAAGTGTGACCTCAGCTT	860
Db	239	AlaGluLysMetAlaGlyCysSerLeuSerPhe	249
RESULT 85			
ID	ADI16668	standard; protein; 324 AA.	
XX	AC	ADI16668;	
XX	XX		
DT	15-APR-2004	(first entry)	
XX			
DE	Human NOVX	protein to treat human pathological conditions	SeqID204.
XX			
KW	human; NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes;		
KW	inflammation; autoimmune disorder; allergy; blood disorder;		
KW	acquired immunodeficiency syndrome; AIDS; obesity; asthma;		
KW	immunoglobulin (IgA) nephropathy; cirrhosis; arthritis;		
KW	Alzheimer's disease; infection; stroke; muscular dystrophy; epilepsy;		
KW	cytostatic; cardiact; antiinflammatory; immunosuppressive; antiallergic;		
KW	haemostatic; anti-HIV; antidiabetic; antiarthritic; hepatotropic;		
KW	antiaesthetic; nephrotropic; antibacterial; virucide; antiparasitic;		
KW	neuroprotective; nootropic; neurogenesis; wound healing; angiogenesis;		
KW	relaxant; anticonvulsant; tissue typing; pharmacogenomic; SNP;		
KW	single nucleotide polymorphism.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200268649-A2.		
XX			
PD	06-SEP-2002.		
XX			
PF	31-JAN-2002; 2002WO-US002785.		
XX			
PR	31-JAN-2001; 2001US-0265395P.		
PR	31-JAN-2001; 2001US-0265412P.		
PR	31-JAN-2001; 2001US-0265514P.		
PR	31-JAN-2001; 2001US-0265517P.		
PR	02-FEB-2001; 2001US-0266406P.		
PR	05-FEB-2001; 2001US-0266767P.		
PR	07-FEB-2001; 2001US-0266975P.		
PR	07-FEB-2001; 2001US-0267057P.		
PR	08-FEB-2001; 2001US-0267459P.		
PR	09-FEB-2001; 2001US-0267823P.		
PR	15-FEB-2001; 2001US-0268974P.		
PR	26-FEB-2001; 2001US-0271664P.		
PR	27-FEB-2001; 2001US-0271839P.		
PR	27-FEB-2001; 2001US-0271855P.		
PR	02-MAR-2001; 2001US-0272788P.		
PR	02-MAR-2001; 2001US-0273046P.		
PR	14-MAR-2001; 2001US-0275925P.		
PR	14-MAR-2001; 2001US-0275947P.		
PR	14-MAR-2001; 2001US-0275950P.		
PR	14-MAR-2001; 2001US-0275989P.		
PR	15-MAR-2001; 2001US-0276448P.		
PR	16-MAR-2001; 2001US-0276450P.		
PR	16-MAR-2001; 2001US-0276397P.		
PR	16-MAR-2001; 2001US-0276768P.		

CC including splice variants) and the nucleic acids (NA) that encode them.
CC Also included are the mature NOVX proteins (and their encoding
CC polynucleotides), a vector comprising NOVX NA, a cell comprising the
CC vector, an antibody that binds immunospecifically to NOVX, determining
CC the presence or amount of NOVX in a sample, determining the presence or
CC amount of NOVX NA in a sample, identifying an agent that binds to NOVX,
CC modulating the activity of NOVX, treating or preventing a NOVX-associated
CC disorder, determining the presence of or predisposition to a disease
CC associated with altered levels of NOVX and treating a pathological state
CC in a mammal comprising administering a polypeptide which is at least 95%
CC identical to NOVX (or fragment). NOVX and NA may be used in the
CC prevention, treatment and diagnosis of diseases associated with
CC inappropriate expression and activity of NOVX (e.g. cancer, diabetes,
CC cardiomyopathy and/or atherosclerosis). The anti-NOVX antibodies and
CC antagonists may also be used to down regulate expression and activity of
CC NOVX. The anti-NOVX antibodies may also be used as diagnostic agents for
CC detecting the presence of NOVX in samples (e.g. by enzyme linked
CC immunosorbent assay (ELISA). The agents and methods may be used in this
CC way to prevent, diagnose and treat cancer, diabetes, cardiomyopathy
CC and/or atherosclerosis. The present sequence represents a NOVX protein.
SQ
XX Sequence 324 AA;

Alignment Scores:
Pred. No.: 7,84e-27 Length: 324
Score: 383.00 Matches: 92
Percent Similarity: 48.1% Conservative: 62
Best Local Similarity: 28.8% Mismatches: 126
Query Match: 8.7% Indels: 40
DB: 8 Gaps: 6

US-10-768-158-1 (1-2419) x ADN42322 (1-324)

QY 21 ATGCGGAGAGCGAGCGCCGAGACCCCGGAGTTCGAG---AGCAAGTAC 77
DB 1 MetAlaLysIleGluLysAsnAlaProThrMetGluLysLysProGluLysPheAsnIle 20
QY 78 TTCGAGTTCATGGCGTGGCGTCCGCCCTTCTGCGCGGGAAGAGAGGAGATCC 137
DB 21 MetGluValAspGlyValProThrLeuIleLeuSerLysGluTrpTrpGluLysValCys 40
QY 138 AACTTCCCGTGGCGCCGAGGAGTG----- 164
DB 41 AsnPheGlnAlaLysProAspAspLeuIleHisAlaSerMetLeuTyrLeuThreGly 60
QY 165 -----TGGATCGTCACCTACCCCAAGTCCGCGCACAGC 197
DB 61 LysLeuProGluGluAspHisGlnAlaTrpLeuGlyAsnTyrProLysSerGlyThrThr 80
QY 198 TTGCTGAGGAGGTGGTCTACTTGTGTAGCCGAGGCGCTGACCCGATGAGATCGGTTG 257
DB 81 TrpMetHisGluIleLeuAspMetIleLeuAsnAspGlyAspValGluLysCysLysArg 100
QY 258 ATGAACATCGAGCAGCAGCTCCCGTCTGGAG-----TACCCA-----CAGCCG 302
DB 101 AlaGlnThrLeuAspArgHisAlaPheLeuGluLeuLysPheProHisLysGluLysPro 120
QY 303 GGCCTGGACATCATCAAGAACTGACCTCTCCCGCTCATCAAGAGCAGCCTGCCCTAC 362
DB 121 AspLeuGluPheValLeuGluMetSerSerProGlnLeuIleLysThrHisLeuProSer 140
QY 363 CGCTTTTCGCTCTGACCTCCCAATGGAGACTCCCAAGGTCTATATATATGCTCGCAAC 422
DB 141 HisLeuIleProProSerIleTrpLysGluAsnCysLysIleValTyrValThrArgAsn 160
QY 423 CCCAGGATCTGGTGGTCTCTATATATCAGTTCACCGCTCTCTCGGACCATGAGTAC 482
DB 161 ProLysAspCysLeuValSerTyrTyrHisPheHisArgMetAlaSerPheMetProAsp 180
QY 483 CGAGGCACCTTTCAGAAATTCGCGGAGGTTTATGAATGATAAGCTGGGCTACGGTCC 542
DB 181 ProGlnAsnLeuGluGluPheTyrGluLysPheMetSerGlyLysValGlyGlySer 200

QY 543 TGGTTTGAGCAGTCGACGAGAGTTCCTGGAGCACCAGCATGGACTCAACGTCGCTTTTCTC 602
DB 201 TrpPheAspHisMetLysGlyTrpTrpAlaAlaLysAspMetHisArgIleLeuTyrLeu 220
QY 603 AAGTATGAGACATGCATCGGACCTGGGACCTGGTGCAGTGGTGGAGCAGCTGGCCAGATTCTCTG 662
DB 221 PheTyrGluAspIleLysLysAsnProLysHisGluIleHisLysValLeuGluPheLeu 240
QY 663 GGGGTGCTCTGTGACAAAGGCCAGCTGGAAGCCCTGACGGAGCAGCTGCCACGAGTGGTG 722
DB 241 GluLysThrTrpSerGlyAspValIleAsnLysIleValHisHisThrSerPheAspVal 260
QY 723 -----GACCACTGCTGCAACGCTGAGGCCCTGCC----- 752
DB 261 MetLysAspAsnProMetAlaAsnHisThrAlaValProAlaHisIlePheAsnHisSer 280
QY 753 -----GTGGCGCGGGAAGAGTTGGGCTGTGGAGGACATCTTCCCGTCTCC 800
DB 281 IleSerLysPheMetArgLysGlyMetProGlyAspTrpLysAsnHisPheThrValAla 300
QY 801 ATGAATGAGAGTTTGACTTGTGTATATAAAGAGATGGGAAAGTGTGACCTCACGTTT 860
DB 301 MetAsnGluAsnPheAspLysHisTyrGluLysLysMetAlaGlySerThrLeuAsnPhe 320
RESULT 87
AAU18710
ID AAU18710 standard; protein; 76 AA.
AC AAU18710;
XX
DT 21-NOV-2001 (first entry)
XX
DE Renal and cardiovascular-associated protein, Seq ID 149.
XX
KW Human; antiinflammatory; neuroprotective; immunomodulator; vulnery;
KW cardiovascular; cystostatic; nephrotropic; antiflaemic; nephritis;
KW immunosuppressive; kidney disorder; renal failure; hypertension;
KW cardiovascular disorder; myocardial infarction; blood disorder; anaemia;
KW blood coagulation disorder; electrolyte imbalance disorder; cancer;
KW hyponatraemia; hyperkalaemia; neoplastic disorder; nephroma;
KW autoimmune disease; inflammatory disease; reproductive system disorder;
KW endocrine disorder; neural activity; neurological disorder;
KW wound healing; respiratory disorder.
XX
OS Homo sapiens.
XX
PN WO200155328-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001359.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227003P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229503P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236803P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239933P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0244647P.
PR 08-NOV-2000; 2000US-0244647P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI
XX WPI; 2001-488787/53.
DR N-PSDB; AAS30231.
XX
PT New polynucleotides and polypeptides, useful for diagnosing, treating,
PT preventing or prognosing e.g. kidney, cardiovascular, blood, electrolyte
XX imbalance or neoplastic disorders, autoimmune diseases, cancers.
PS Claim 1; SEQ ID NO 149; 506pp; English.
XX
CC The invention relates to novel nucleic acids and polypeptides useful for
CC diagnosing, treating, preventing and/or prognosing disorders related to
CC these polypeptides. The polynucleotides are especially useful in the
CC diagnosis, prognosis, prevention and/or treatment of diseases which
CC include kidney disorders (e.g. renal failure or nephritis),
CC cardiovascular disorders (e.g. hypertension or myocardial infarction),
CC blood disorders (e.g. anaemia or blood coagulation disorders),
CC electrolyte imbalance disorders (e.g. hyponatraemia or hyperkalaemia),
CC neoplastic disorders (e.g. nephroma or renal cell cancer), autoimmune
CC diseases, cancers, inflammatory diseases, reproductive system disorders,
CC endocrine disorders, neural activity and neurological disorders, wound
CC healing and respiratory disorders. AAU18644-AAU18715 represent the novel
CC human renal and cardiovascular-associated amino acid sequences of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
XX from WIPO at: ftp.wipo.int/pub/published_pct_sequences

PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250391P.
PR 01-DEC-2000; 2000US-0251160P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251473P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-541565/60.
XX N-PSDB; ABA11230.
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
XX useful for preventing, diagnosing and/or treating nervous system cancers
XX and metastases.
XX Claim 11; SEQ ID NO 3561; 1701pp + Sequence Listing; English.
XX The invention relates to novel genes (ABA11004-ABA21534) and proteins
XX (ABA114678-ABA18001) useful for preventing, treating or ameliorating
XX medical conditions e.g. by protein or gene therapy. The genes are
XX isolated from a range of human tissues disclosed in the specification.
XX The nucleic acids, proteins, antibodies and (ant)agonists are useful in
XX the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
XX ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
XX breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
XX disorders e.g. Addison's disease, allergies, autoimmune haemolytic
XX anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
XX multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
XX cardiovascular disorders such as myocardial ischaemia; (d) wound healing
XX ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
XX infectious diseases such as viral, bacterial, fungal and parasitic
XX infections. Note: The sequence data for this patent did not form part of

CC the printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences
SQ Sequence 76 AA;
Alignment Scores:
Pred. No.: 4.52e-27 Length: 76
Score: 382.00 Matches: 71
Percent Similarity: 97.3% Conservative: 1
Best Local Similarity: 95.9% Mismatches: 2
Query Match: 8.7% Indels: 0
DB: Gaps: 0
US-10-768-158-1 (1-2419) x ABB14904 (1-76)
Qy 288 GAGTACCCACAGCGCGCTGGACATCATCAAGGACTGACCTCTCCCGCTCATCAAG 347
Db ::::||||| 1 LysTyrProGlnProGlyLeuAspIleile***GluLeuThrSerProArgLeuLys 20
Qy 348 AGCCACCTGGCCTACCGCTTTCTGCCCTCTGACCTCCACAAATGGAGACTCCAAAGGTCTATC 407
Db ||||| 21 SerHisLeuProTyrArgPheLeuProSerAspLeuHisAsnGlyAspSerLysValle 40
Qy 408 TATATGGCTCGCAACCCAAAGGATCTGGTGTGCTTATTATCAGTTCCACCGCTCTCTG 467
Db ||||| 41 TyrMetAlaArgAsnProLysAspLeuValValSerTyrTyrGlnPheHisArgSerLeu 60
Qy 468 CGGACCATGAGCTACCGAGGACCTTTCAAGAAATTCCTGCCGG 509
Db ||||| 61 ArgThrMetSerTyrArgGly***PheGlnGluPheCysArg 74
RESULT 89
ABU97325
ID ABU97325 standard; protein; 76 AA.
XX AC ABU97325;
XX DT 30-JUL-2003 (first entry)
XX DE Human polypeptide #67.
XX KW Human; kidney disorder; cardiovascular disorder; arrhythmia;
KW glomerulonephritis; urinary tract infection; chronic nephritis; anaemia;
KW carcinoid heart disease; endocarditis; blood disorder; thrombosis;
KW haemoglobin abnormality; electrolyte imbalance; neoplastic disorder;
KW cancer; respiratory disorder; acute rhinitis; sinusitis; pharyngitis;
KW neurological disorder; Alzheimer's disease; Parkinson's disease;
KW Huntington's disease.
XX OS Homo sapiens.
XX PN US2003013649-A1.
XX PD 16-JAN-2003.
XX PF 21-NOV-2001; 2001US-00989442.
XX PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241828P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.

PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-00764863.

(ROSE/) ROSEN C A.
(RUBE/) RUBEN S M.
(BARA/) BARASH S C.

Rosen CA, Ruben SM, Barash SC;

WPI; 2003-416807/39.
N-PSDB; ACD01486.

New nucleic acids and polypeptides, useful for diagnosing, prognosing, preventing and/or treating e.g. kidney, cardiovascular, blood, electrolyte imbalance, neoplastic, respiratory, or neurological diseases or disorders.

Claim 11; Page 304; 363pp; English.

The invention relates to isolated nucleic acids encoding novel polypeptides. The nucleic acids are useful for chromosome mapping, for radiation hybrid mapping, for detection of cancer, in gene therapy, for identifying individuals from minute biological samples, as an alternative to restriction fragment length polymorphism (RFLP) analysis, in forensic biology and as hybridisation probes for differential identification of tissues or cell types present in a biological sample. Compositions comprising the polynucleotides, polypeptides and antibodies specific for

CC the polypeptides may be used in the diagnosis, prognosis, prevention
CC and/or treatment of kidney disorders (e.g. glomerulonephritis, urinary
CC tract infections, chronic nephritis), cardiovascular disorders (e.g.
CC arrhythmias, carotid heart disease, endocarditis), blood disorders
CC (e.g. thrombosis, anaemia, haemoglobin abnormalities), electrolyte
CC imbalance, neoplastic disorders (e.g. cancers), respiratory disorders
CC (e.g. acute rhinitis, sinusitis, pharyngitis) and neurological disorders
CC (e.g. Alzheimer's disease, Parkinson's disease, Huntington's disease).
CC Sequences ABU97259-ABU97330 represent human polypeptides of the invention
XX

SQ Sequence 76 AA;

Alignment Scores:
Pred. No.: 4, 52e-27 Length: 76
Score: 382.00 Matches: 71
Percent Similarity: 97.3% Conservative: 1
Best Local Similarity: 95.9% Mismatches: 2
Query Match: 8.7% Indels: 0
DB: 6 Gaps: 0

US-10-768-158-1 (1-2419) x ABU97325 (1-76)

QY 288 GAGTACCACAGCGCGCTGACATATCAAGAACTGACTCTCCCGCTCATCAAG 347
Db 1 LysTyrProGlnProGlyLeuAspIleile**GluLeuThrSerProArgLeuIlelys 20
QY 348 AGCCACTGCGCTACCGCTTCTGCGCTCTGACCTCCACATGGAGACTCCAAAGTGCATC 407
Db 21 SerHisLeuProTyrArgPheLeuProSerAspLeuHisAsnGlyAspSerIysValile 40
QY 408 TATATGGCTCGCAACCCCAAGGATCTGGTGTCTTATTATCAGTTCACCGCTCTCTG 467
Db 41 TyrMetAlaArgAsnProIysAspLeuValSerIysTyrGlnPheHisArgSerLeu 60
QY 468 CGGACCATGAGTACCGAGGACCTTTCAAGAATTCTCCCGG 509
Db 61 ArgThrMetSerTyrArgGly***PheGlnGluPheCysArg 74

RESULT 90
ABM84007
ID ABM84007 standard; protein; 254 AA.

XX AC ABM84007;

XX DT 18-NOV-2004 (first entry)

XX DE Human diagnostic and therapeutic pprotein SEQ ID NO:4256.

XX KW gene therapy; human diagnostic and therapeutic polynucleotide; dithp.

XX OS Homo sapiens.

XX PN WO2004023973-A2.

XX PD 25-MAR-2004.

XX PF 12-SEP-2003; 2003WO-US028227.

XX PR 12-SEP-2002; 2002US-0410259P.

XX PR 12-SEP-2002; 2002US-0410260P.

XX PA (INCY-) INCYTE CORP.

XX PI Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
PI Harthshorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;
PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
PI Patury S, Shi X, Suarez CJ;
XX WPI; 2004-329368/30.

DR N-PSDB; ACN42659.
XX New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT in diagnosing a condition, disease or disorder associated with human
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT in gene mapping.
XX Claim 27; Page; 190pp; English.

XX The invention relates to novel diagnostic and therapeutic polynucleotides
CC selected from one of the 2722 sequences defined in the specification. A
CC polynucleotide of the invention may have a use in gene therapy. The human
CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
CC used to diagnose a particular condition, disease or disorder associated
CC with human molecules, e.g. cell proliferative disorders
CC autoimmune/inflammatory disorder, developmental disorder, endocrine
CC disorder, neurological disorders, gastrointestinal disorders, or
CC infections caused by virus, bacteria, fungi or parasite. The dithp
CC molecules may also be used in genetic mapping, in identifying individuals
CC from minute biological samples, in detecting single nucleotide
CC polymorphisms, as molecular weight markers, and for somatic or germline
CC gene therapy. The present sequence represents a dithp protein of the
CC invention. Note: The sequence data for this patent is not represented in
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm

XX SQ Sequence 254 AA;

Alignment Scores:
Pred. No.: 2, 87e-26 Length: 254
Score: 376.50 Matches: 83
Percent Similarity: 47.6% Conservative: 48
Best Local Similarity: 30.2% Mismatches: 85
Query Match: 8.6% Indels: 59
DB: 8 Gaps: 4

US-10-768-158-1 (1-2419) x ABM84007 (1-254)

QY 90 GCGTGGCGGTGCGCGCTTCTGCGCGGAGATGAGGAGATCGCAACTCCCGTG 149
Db 17 GlyValProLeuIleLysTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGluAla 36
QY 150 CGGCCACGACGCTGGATCGTCACCTACCCCAAGTCCCGCACAGCTTCTCGACGAG 209
Db 37 ArgProAspLeuLeuLeuAsnThrTyrProLysSerGlyThrTrpValSerGln 56
QY 210 GTGCTCTACTTGTGAGCCAGGCGCTGACCCCGATGAGATCGGCTTGATGAACATGCAC 269
Db 57 IleLeuAspMetIleTyrGlnGlyGlyAspLeuGluLysCysAsnArgAlaProIleTyr 76
QY 270 GAGCAGCTCCCGTCTCGGAGTACCCACGCGCGGCTGGACATCATCAAGAACTGACC 329
Db 77 ValArgValProPheLeuGlu----- 83
QY 330 TCTCCCGCTCATCAAGAGCCACCTGCGCTTCTGCGCTTCTGACCTCCACAAT 389
Db 83 ----- 83
QY 390 GGAGACTCCAAGGTCACTATATGCTCGCAACCCCAAGGATCTGGTGTCTTATTAT 449
Db 84 -----ValValTyrValAlaArgAsnProLysAspValAlaValSerTyrTyr 99
QY 450 CAGTTCACCGCTCTCTGCGGACCATGAGTACCGAGGACCTTTCAGAATTTCTGCGG 509
Db 100 HisPheHisArgMetGluLeuAlaHisProGluProGlyThrTrpAspSerPheLeuGlu 119
QY 510 AGGTTTATGAATATAGCTGGCTGCGCTCTCGTGTGTTGAGCACGCTGACGAGTCTTGG 569
Db 120 LysPheMetAlaGlyGluValSerTyrGlySerTrpTyrGlnHisValGlnGluTrpTrp 139
QY 570 GAGCACCAGCATGAGTCTCGAAGCTGCTTTTCTCAAGTATGAAGACATCATCGGACCTG 629
Db 140 GluLeuSerArgThrHisProValLeuTyrLeuPheTyrGluAspMetLysGluAsnPro 159

Qy	816	GACTTGGTGATATAACAGAGATGGAAAGTGTGACCTCAGCTTT	860
Db	236	AspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe	250
RESULT 92			
ABM84006			
ID	ABM84006	standard; protein; 254 AA.	
XX	AC	ABM84006;	
XX	XX	18-NOV-2004 (first entry)	
XX	DE	Human diagnostic and therapeutic pprotein SEQ ID NO:4255.	
XX	KW	gene therapy; human diagnostic and therapeutic polynucleotide; dithp.	
XX	OS	Homo sapiens.	
XX	PN	WO2004023973-A2.	
XX	PD	25-MAR-2004.	
XX	PF	12-SEP-2003; 2003WO-US028227.	
XX	PR	12-SEP-2002; 2002US-0410259P.	
XX	PR	12-SEP-2002; 2002US-0410260P.	
XX	PA	(INCY-) INCYTE CORP.	
XX	PI	Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F; Hatthesorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV; Mooney EM, Deleageane AM, Panesar IS, Banville SC, Reddy TP; Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH; Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve Lu; Lagace RE, Spito PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES; Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D; Patury S, Shi X, Suarez CJ;	
XX	PI	WPI; 2004-329368/30.	
DR	DR	N-PSDB; ACN42658.	
XX	PT	New diagnostic and therapeutic polynucleotides and polypeptides, useful in diagnosing a condition, disease or disorder associated with human molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or in gene mapping.	
XX	PS	Claim 27; Page; 190pp; English.	
XX	CC	The invention relates to novel diagnostic and therapeutic polynucleotides selected from one of the 2722 sequences defined in the specification. A polynucleotide of the invention may have a use in gene therapy. The human diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be used to diagnose a particular condition, disease or disorder associated with human molecules, e.g. cell proliferative disorders, autoimmune/inflammatory disorder, developmental disorder, endocrine disorder, neurological disorders, gastrointestinal disorders, or infections caused by virus, bacteria, fungi or parasite. The dithp molecules may also be used in genetic mapping, in identifying individuals from minute biological samples, in detecting single nucleotide polymorphisms, as molecular weight markers, and for somatic or germline gene therapy. The present sequence represents a dithp protein of the invention. Note: The sequence data for this patent is not represented in the printed specification, but was obtained in electronic format directly from WIPO at www.wipo.int/pct/en/sequences/listing.htm	
XX	SQ	Sequence 254 AA;	
Alignment Scores:			
Prod. No.:	2.87e-26	Length:	254
Score:	376.50	Matches:	83
Percent Similarity:	47.6%	Conservative:	48
Best Local Similarity:	30.2%	Mismatches:	85
Query Match:	8.6%	Indels:	59
DB:	8	Gaps:	4
US-10-768-158-1 (1-2419) x ABM84006 (1-254)			
Qy	90	GGCGTGGCTGCGCGCTTCTGCGCGGGAAGATGGAGGAGATCGCAACTTCCCGGTG	149
Db	17	GlyValProLeuIleLysTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla	36
Qy	150	CGGCCAGCAGCGTGTGGATCGTCACCTACCCCAAGTCCGGCAGCAGCTTGTGCGAGG	209
Db	37	ArgProAspAspLeuLeuIleAsnThrTyrProLysSerGlyThrTrpValSerGln	56
Qy	210	GTGCTCTACTTGTGAGCCAGGCGCTGACCCCGATGAGATCGCTTGTGATGAACATCGAC	269
Db	57	IleuAspMetIleTyrGlnGlyGlyAspLeuGluLysCysAsnArgAlaProIleTyr	76
Qy	270	GAGCAGCTCCCGGTCTCTGGAGTACCCACAGCCGGCGCTGGACATCATCAAGGAAC	329
Db	77	ValArgValProPheLeuGlu	83
Qy	330	TCGCCCGCTCATCAAGAGCCACCTGCCCTTCTGCCCTCTGACCTCCACAAT	389
Db	83		83
Qy	390	GGAGACTCCAAGGTCTATATGCTCGCAACCCCAAGGATCTGGTGTGCTTTATTAT	449
Db	84	-----ValValTyrValAlaAlaArgAsnProLysAspValAlaValSerTyr	99
Qy	450	CAGTTCCACCGCTCTCTGCGGACCATGAGTACCGAGGACCTTTCGAAGAAATTC	509
Db	100	HisPheHisArgMetGluLysAlaHisProGluProGlyThrTrpAspSerPheLeuGlu	119
Qy	510	AGTTTATGAATAGTACGCTGCGCTCTGCTGTTGAGCAGCGTGGAGGATTCGG	569
Db	120	LysPheMetAlaGlyGluValSerTyrGlySerTyrGlnHisValGlnGluTrpTrp	139
Qy	570	GAGCAGCGCTGAGTACGAGCTGCTTTCTCAAGTATCAAGACATGACATCGGACCTG	629
Db	140	GluLeuSerArgThrHisProValLeuTyrLeuPheTyrGluAspMetLysGluAsnPro	159
Qy	630	GTGACGATGTGGAGCAGCTGGCCAGATTCCTGGGGTGTCTGTGACAAGGCCAGCTG	689
Db	160	LysArgGluIleGlnLysIleLeuGluPheValGlyArgSerLeuProGluGluThrMet	179
Qy	690	GAAGCCCTGACGAGCACTGC-----	710
Db	180	AspPheMetValGlnHisThrSerPheLysGluMetLysLysAsnProMetThrAsnTyr	199
Qy	711	-----CACACAGCTGGTGGACCACTGTGCAACGCTGAGGCCCTGCCC---GTG	755
Db	200	ThrThrValProGlnGluLeuMetAspHisSerIleSer-----ProPheMet	215
Qy	756	GGCCGGGGAAGATTGGCTGTGGAGGACATCTTACCTCTCCATGAATGAAGATTT	815
Db	216	ArgLysGlyMetAlaGlyAspTrpLysThrThrPheThrValAlaGlnAsnGluArgPhe	235
Qy	816	GACTTGGTGATATAACAGAGATGGGAAGTGTGACCTCAGCTT	860
Db	236	AspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe	250
RESULT 93			
ABM82619			
ID	ABM82619	standard; protein; 254 AA.	
XX	AC	ABM82619;	
XX	XX	18-NOV-2004 (first entry)	
XX	DE	Human diagnostic and therapeutic pprotein SEQ ID NO:2868.	
XX	KW	gene therapy; human diagnostic and therapeutic polynucleotide; dithp.	
XX	OS	Homo sapiens.	

PI Legace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirtson ES;
PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
PI Patury S, Shi X, Suarez CJ;
XX
DR N-PSDB; ACN41269.
XX
XX
PT New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT in diagnosing a condition, disease or disorder associated with human
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT in gene mapping.
XX
XX
PS Claim 27; Page; 190pp; English.
XX
XX The invention relates to novel diagnostic and therapeutic polynucleotides
CC selected from one of the 2722 sequences defined in the specification. A
CC polynucleotide of the invention may have a use in gene therapy. The human
CC polynucleotide and therapeutic polynucleotides (dithp) or polypeptides may be
CC used to diagnose a particular condition, disease or disorder associated
CC with human molecules, e.g. cell proliferative disorders,
CC autoimmune/inflammatory disorders, developmental disorder, endocrine
CC disorder, neurological disorders, gastrointestinal disorders, or
CC infections caused by virus, bacteria, fungi or parasite. The dithp
CC molecules may also be used in genetic mapping, in identifying individuals
CC from minute biological samples, in detecting single nucleotide
CC polymorphisms, as molecular weight markers, and for somatic or germline
CC gene therapy. The present sequence represents a dithp protein of the
CC invention. Note: The sequence data for this patent is not represented in
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX
XX
SQ Sequence 254 AA;

Alignment Scores:

Pred. No.: 2.87e-26 Length: 254
Score: 376.50 Matches: 83
Percent Similarity: 47.6% Conservative: 48
Best Local Similarity: 30.2% Mismatches: 85
Query Match: 8.6% Indels: 59
DB: 8 Gaps: 4

US-10-768-158-1 (1-2419) x ABM82617 (1-254)

Qy 90 GGCGTGGCGGCTGCGCCCTTCTGCGCGGAGATGGAGAGATCGCCAACTCCCGTG 149
Dy 17 GlyValProLeuIleLysTyPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla 36
Qy 150 CGGCCACGCGAGTGTGGATGTCACCTACCCCAAGTCCGACAGCTTCTGCGAGGAG 209
Dy 37 ArgProAspLeuLeuIleAsnThrTyProLysSerGlyThrTrpValSerGln 56
Qy 210 GTGGTCTACTTGGTGGAGCGGCGCTGACCCCGATGATGCGGCTGATGAACATGAC 269
Dy 57 IleLeuAspMetIleTyGlnGlyAspLeuGluLysCysAsnArgAlaProIleTy 76
Qy 270 GAGCAGCTCCGGTCTCGAGTACCCACAGCGGCGCTGGACATCATCAAGGAAGTACC 329
Dy 77 ValArgValProPheLeuGlu----- 83
Qy 330 TCTCCCGCCCTCATCAAGAGCCACCTGCGCTTCTGCGCTTCTGCGCTTCTGCGCAAT 389
Dy 83 ----- 83
Qy 390 GGAGACTCCAAGTCTATATGCTCGCAACCCCAAGATCTGGTGGTCTTATAT 449
Dy 84 -----ValValTyValAlaArgAsnProLysAspValAlaValSerTyTy 99
Qy 450 CAGTCCACCGCTCTCTCGGAGCATTGAGTACCGAGGACCTTCAAGAATTCTGCGG 509
Dy 100 HisPheHisArgMetGluLysAlaHisProGluProGlyThrTrpAspSerPheLeuGlu 119
Qy 510 AGGTTTATGAATGATGAGTGGCTGACCGCTCTGTTGAGCAGCTGCAGGAGTCTGG 569
Dy ----- 569

Dy 120 LysPheMetAlaGlyGluValSerTyGlySerTrpTyGlnHisValGlnGluTrpTrp 139
Qy 570 GAGCACCCTGACTCGAAGCTCTTTTCTCAAGTATGAAGACATCGGACCTG 629
Dy 140 GluLeuSerArgThrHisProValLeuTyLeuPheTyGluAspMetLysGluAsnPro 159
Qy 630 GTGACGATGTGGAGCAGCTGGCCAGATTCTCGGGGTGCTCTGACAAAGCCAGCTG 689
Dy 160 LysArgGluIleGlnLysIleLeuGluPheValGlyArgSerLeuProGluGluThrMet 179
Qy 690 GAAGCCCTGACGGAGCAGCTGC----- 710
Dy 180 AspPheMetValGlnHisThrSerPheLysGluMetLysLysAsnProMetThrAsnTy 199
Qy 711 -----CACACGCTGGTGACAGCTGCTGCAACGCTGAGGCCCTGCC---GTG 755
Dy 200 ThrThrValProGlnGluLeuMetAspHisSerIleSer-----ProPheMet 215
Qy 756 GGCCGGGGAAGAGTTGGCTGTGGAAGGACATCTTACACCGTCTCCATGAATGAGAAGTTT 815
Dy 216 ArgLysGlyMetAlaGlyAspTrpLysThrPheThrValAlaGlnAsnGluArgPhe 235
Qy 816 GACTTGTGTATAACAGAGATGGAAAGTGTGACCTCACGTTT 860
Dy 236 AspAlaAspTyAlaGluLysMetAlaGlyCysSerLeuSerPhe 250
RESULT 95
ABM82618
ID ABM82618 standard; protein; 254 AA.
XX
AC ABM82618;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human diagnostic and therapeutic pprotein SEQ ID NO:2867.
XX
KW gene therapy; human diagnostic and therapeutic polynucleotide; dithp.
XX
OS Homo sapiens.
XX
PN WO2004023973-A2.
XX
PD 25-MAR-2004.
XX
PF 12-SEP-2003; 2003WO-US028227.
XX
PR 12-SEP-2002; 2002US-0410259P.
XX
PR 12-SEP-2002; 2002US-0410260P.
XX
PA (INCY-) INCYTE CORP.
XX
PI Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
PI Harthorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
PI Mooney EM, Deleagane AM, Panesar IS, Banville SC, Reddy TP;
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin BH;
PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
PI Legace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirtson ES;
PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
PI Patury S, Shi X, Suarez CJ;
XX
XX WPI; 2004-329368/30.
XX N-PSDB; ACN41270.
XX
XX New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT in diagnosing a condition, disease or disorder associated with human
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT in gene mapping.
XX
PS Claim 27; Page; 190pp; English.
XX
XX The invention relates to novel diagnostic and therapeutic polynucleotides
CC selected from one of the 2722 sequences defined in the specification. A
CC polynucleotide of the invention may have a use in gene therapy. The human
CC polynucleotide of the invention may have a use in gene therapy. The human

diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be used to diagnose a particular condition, disease or disorder associated with human molecules, e.g. cell proliferative disorders, autoimmune/inflammatory disorder, developmental disorder, endocrine disorder, neurological disorders, gastrointestinal disorders, or infections caused by virus, bacteria, fungi or parasite. The dithp molecules may also be used in genetic mapping, in identifying individuals from minute biological samples, in detecting single nucleotide polymorphisms, as molecular weight markers, and for somatic or germline gene therapy. The present sequence represents a dithp protein of the invention. Note: The sequence data for this patent is not represented in the printed specification, but was obtained in electronic format directly from WIPO at www.wipo.int/pct/en/sequences/listing.htm

Sequence 254 AA;

Alignment Scores:	
Pred. No.:	2, 87e-26
Score:	376.50
Percent Similarity:	47.6%
Best Local Similarity:	30.2%
Query Match:	8.6%
DB:	8
Length:	254
Matches:	83
Conservative:	48
Mismatches:	85
Indels:	59
Gaps:	4

US-10-768-158-1 (1-2419) x ABM82618 (1-254)

Qy	90	GGCGTGGCGGTCTGCGCCCTTCTGCGCGGGAGATGGAGAGATCCCAACTTCCTCCGGT	149
Db	17	 ::: GlyValProLeuIleLysTyPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla	36
Qy	150	CGGCGCCAGCGAGTGTGATCGTCACTACCCCAAGTCCGGCACACAGCTGTCTGCAGGAG	209
Db	37	 ::: ArgProAspAspLeuLeuIleAsnThrTyProLysSerGlyThrTrpValSerGln	56
Qy	210	GTGCTTACTTGGTGAGCGCGCTCACCCGATGAGTCGGCTTGATGAACATCGAC	269
Db	57	 ::: IleLeuAspMetIleTyGlnGlyGlyAspLeuGluLysCysAsnArgAlaProIleTy	76
Qy	270	GAGCAGCTCCCGTCTCTGGAGTACCACAGCGGGCGCTGGACATCATCAAGGAAC	329
Db	77	 ::: ValArgValProPheLeuGlu	83
Qy	330	TCTCCCGCCTCATCAAGAGCCACTGCGCTTACCGCTTCTGCGCTCTGACCTCCACAAT	389
Db	83	-----	83
Qy	390	GGAGACTCCAAGGTCACTTATATGCTCGCAACCCCAAGGATCTGGTGTCTTATTAT	449
Db	84	 ::: ValValTyValAlaArgAsnProLysAspValAlaValSerTyTrp	99
Qy	450	CAGTTCCACCGCTCTCTCGCGACCATGAGTACCAGGACCTTTCAAGAAATCTTCGCGG	509
Db	100	 ::: HisPheHisArgMetGluLysAlaHisProGluProGlyThrTrpAspSerPheLeuGlu	119
Qy	510	AGTTTATGAATGATAGCTGGGTACGGCTCTGGTTTGAGCACGTGCAGGAGTTCTGG	569
Db	120	 ::: LysPheMetAlaGlyGluValSerTyGlySerTrpTyrGlnHisValGlnGluTrpTrp	139
Qy	570	GAGCACCGCATGGACTGAAACGTGCTTTTCTCAAGTATGAAGACATGATCGGACCTG	629
Db	140	 ::: GluLeuSerArgThrHisProValLeuTyLeuPheTyGluAspMetLysGluAsnPro	159
Qy	630	GTGACGATGTGGAGCAGCTGGCCAGATTCTCTGGGGGTCTCTGTGACAAGSCCAGCTG	689
Db	160	 ::: LysArgGluIleGlnLysIleLeuGluPheValGlyArgSerLeuProGluGluThrMet	179
Qy	690	GAAGCCCTGACGGACACTGC	710
Db	180	 ::: AspPheMetValGlnHisThrSerPheLysGluMetLysLysAsnProMetThrAsnTy	199
Qy	711	-----	711
Db	200	 ::: ThrThrValProGlnGluLeuMetAspHisSerIleSer-----ProPheMet	215
Qy	755	-----	755

Qy	756	GGCCGGGAGAGAGTTGGCTGTGGAAGGACATCTTCACCGTCTCCATGAATGAGAAGTTT	815
		:::	
Db	216	ArgLysGlyMetAlaGlyAspTrpLysThrThrPheThrValAlaGlnAsnGluArgPhe	235
		:::	
Qy	816	GACTTGGTGTATAAACAGAAAGTGGGAAAAGTGTGACCTCACGTTT	860
		:::	
Db	236	AspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe	250
		:::	
RESULT_96			
ABM82620			
ID	ABM82620 standard; protein; 254 AA.		
XX			
XX	ABM82620;		
XX			
DT	18-NOV-2004 (first entry)		
XX			
XX	Human diagnostic and therapeutic pprotein SEQ ID NO:2869.		
DE			
XX			
KW	gene therapy; human diagnostic and therapeutic polynucleotide; dithp.		
XX			
XX	Homo sapiens.		
XX			
PN	WO2004023973-A2.		
XX			
PD	25-MAR-2004.		
XX			
PF	12-SEP-2003; 2003WO-US028227.		
XX			
PR	12-SEP-2002; 2002US-0410259P.		
PR	12-SEP-2002; 2002US-0410260P.		
XX			
PA	(INCY-) INCYTE CORP.		

Sequence 254 AA;

Alignment Scores:


```
Pred. No.: 2.87e-26 Length: 254
Score: 376.50 Matches: 83
Percent Similarity: 47.6% Conservatives: 48
Best Local Similarity: 30.2% Mismatches: 85
Query Match: 8.6% Indels: 59
DB: 8 Gaps: 4

US-10-768-158-1 (1-2419) x ABM82620 (1-254)
QY 90 GCGGTGCGGCTGCGCCCTTCTGCGCGGGAAGATGGAGGAGATCGCCAACTTCCCGGTG 149
DB 17 GlyValProLeuIleLysTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla 36
QY 150 CGGCCACGACGCTGATCGTACCTACCCAGTCGCGCACGCTTCTCGCAGGAG 209
DB 37 ArgProAspAspLeuLeuIleAsnThrTyrProLysSerGlyThrTrpValSerGln 56
QY 210 GTGGTCTACTTGGTGAGCAGCGGCTGACCCGATGAGATCGGCTTGATGAACATCGAC 269
DB 57 IleLeuAspMetIleTyrGlnGlyGlyAspLeuGluLysCysAsnArgAlaProIleTyr 76
QY 270 GAGCAGCTCCCGGTCTCGAGTACCCAGCGCGGCTGGACATCATCAAGAACTGACC 329
DB 77 ValArgValProPheLeuGlu----- 83
QY 330 TCTCCCCGCTCATCAAGAGCACCTGCCCTTACCGCTTCTGCGCTTCGACCTCCACAT 389
DB 83 ----- 83
QY 390 GGAGACTCAAGGTCTATATGCTCGCAACCCCAAGGATGCTGGTGTCTTATTAT 449
DB 84 -----ValValTyrValAlaArgAsnProLysAspValAlaValSerTyrTyr 99
QY 450 CAGTTCCACCGCTCTCTCGGACCATGAGTACCGAGCACCTTTCAAGAATTTGCGCG 509
DB 100 HisPheHisArgMetGluLysAlaHisProGluProGlyThrTrpAspSerPheLeuGlu 119
QY 510 AGGTTTATGATGATGAAGCTGGCTACGGCTCTGTTTGGACGCTGACGAGTCTGG 569
DB 120 LysPheMetAlaGlyGluValSerTyrGlySerTyrTyrGlnHisValGlnGluTrpTrp 139
QY 570 GAGCACCAGTGGACTCGAACGTGCTTTTCTCAAGTATGAAGACATCATCGGACCTG 629
DB 140 GluLeuSerArgThrHisProValLeuTyrLeuPheTyrGluAspMetLysGluAsnPro 159
QY 630 GTGACGATGGTGGACAGCTGCCAGATTCCTGGGGGTGTCTCTGTCACAAAGCCAGCT 689
DB 160 LysArgGluIleGlnLysIleLeuGluPheValGlyArgSerLeuProGluGluThrMet 179
QY 690 GAAGCCCTGACGGAGCAGCTGC----- 710
DB 180 AspPheMetValGlnHisThrSerPheLysGluMetLysLysAsnProMetThrAsnTyr 199
QY 711 -----CACCAGCTGTGACAGTGTCTCAACGCTGAGCGCCTGCC-- -GTG 755
DB 200 ThrThrValProGlnGluLeuMetAspHisSerIleSer-----ProPheMet 215
QY 756 GCGCGGGAAGAGTTGGGTGTGGGAAGACATCTTCCCGCTCTCCATGATGAGAAGTTT 815
DB 216 ArgLysGlyMetAlaGlyAspTrpLysThrThrPheThrValAlaGlnAsnGluArgPhe 235
QY 816 GACTTGGTGTATAACACAGAAGATGGAAAGTGTGACCTCACGTTT 860
DB 236 AspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe 250

RESULT 97
ADG76311
ID ADG76311 standard; protein; 259 AA.
XX
AC ADG76311;
XX
DT 11-MAR-2004 (first entry)
XX
```

```
DE Human Incyte enzymatic protein (SeqID 3).
XX enzyme; human; biogenesis; biodegradation; autoimmune; inflammatory;
KW infectious; neurological; cardiovascular; metabolic; cell proliferative;
KW cancer; cytostatic; immunomodulator.
XX
OS Homo sapiens.
XX
PN WO2003083082-A2.
XX
PD 09-OCT-2003.
XX
PF 26-MAR-2003; 2003WO-US009772.
XX
PR 29-MAR-2002; 2002US-0368721P.
PR 29-MAR-2002; 2002US-0368799P.
XX
PA (INCY-) INCYTE CORP.
XX
PI Tran UK, Kable AE, Ramkumar J, Ison CH, Richardson TW, Lee SY,
PI Khare R, Marquis JP, Swarnakar A, Chawla NK, Elliott VS,
PI Emerling BM, Becha SD, Hafalia AJA, Li JX, Griffin JA, Hawkins PR,
PI Jin P, Chien D, Jiang X, Jackson AA, Mason PM, Bhatia UG,
PI Burtrill JD, Lee S, Blake JJ, Ho A, Zheng W, Gao J;
XX
DR WPI; 2003-804054/75.
DR N-PSDB; ADG76366.
XX
PT New human enzymes (ENZM) polypeptide, useful for preparing a composition
PT for treating a disease associated with decreased expression or
PT overexpression of ENZM e.g. cancer.
XX
PS Claim 58; SEQ ID NO 3; 385pp; English.
XX
CC This invention relates to novel isolated polynucleotides, and the enzymes
CC encoded thereof, as well as any biologically active or immunogenic
CC fragments. Specifically, it refers to human enzymes involved in the
CC cellular processes of biogenesis and biodegradation, for example
CC oxidoreductases, hydrolases and isomerases. The present invention
CC describes the use of these nucleic acids and proteins (including single
CC nucleotide polymorphisms) in the diagnosis, treatment and prevention of
CC various disorders such as autoimmune/ inflammatory, infectious,
CC neurological, cardiovascular, metabolic and cell proliferative such as
CC cancer. Furthermore, the polypeptide is useful for preparing a
CC composition for treating a disease associated with decreased expression
CC or overexpression of the functional enzyme (ENZM) e.g. immune disorders.
CC Accordingly, the compositions exhibit cytostatic and immunomodulator
CC activities. This polypeptide sequence is a human Incyte protein, an
CC enzyme of the invention.
XX
SQ Sequence 259 AA;

Alignment Scores:
Pred. No.: 5.04e-25 Length: 259
Score: 363.50 Matches: 83
Percent Similarity: 48.6% Conservatives: 53
Best Local Similarity: 29.8% Mismatches: 91
Query Match: 8.3% Indels: 53
DB: 7 Gaps: 5

US-10-768-158-1 (1-2419) x ADG76311 (1-259)
QY 72 AAGTACTTCGAGTTCATGCTGGGTGCGGCTTCTCGCGGGAAGATGGAGGAG 131
DB 12 LysLeuLysGluValGluGlyThrLeuLeuGlnProAlaThrValAspAsnTrpSerGln 31
QY 132 ATCGCCCACTTCCCGTGGCGGCGGACGAGCTGTGGATCGTCACTACCCCAAGTCCGCG 191
DB 32 IleGlnSerPheGluAlaLysProAspAspLeuIleCysThrTyrProLysAlaGly 51
QY 192 ACCAGCTTGTGCAGGAGGTGCTTACTTGTGTAGCCAGCGGCTGACCCCGATGAGATC 251
DB 52 ThrThrTrpIleGlnIleValAspMetIleGluGlnAsnGlyAspValGluLysCys 71
```



```
Db 216 ProhennilePheThrSerTyrGluArgMetIysGlyGlnLeuGlyGlnValIleSer 235
Qy 645 CAGCTGCCAGATTCTCTGGGGTGTCTGTGTACAAAGCCCGAGCTGGAGCCCTGACGGAG 704
Db 236 GluValAlaGlnPheLeuGluArgSerValSerGlnGlnMetGlnMetGlnArg 255
Qy 705 CACTGC-----CACCAGCTGGTGACAG---TGCTGCAAC----- 737
Db 256 HisLeuSerPheGluSerMetArgAsnProAlaCysAsnHisValIysGluPheGlu 275
Qy 738 -----GCTGAGGCCCTGCCCGTGGCGCGGGGAAGA 767
Db 276 SerMetIysAlaAlaGlyArgGluValGluGluPheArgPheValArgGlyVal 295
Qy 768 GTTGGCTGTGGAGGACATCTTCACCGTCTCCATGAATGAGAAAGTTTGACTTG 821
Db 296 ValGlySerHisIysAspGluLeuThrAlaAspIleIleArgGluPheAspLeu 313
RESULT 99
AB862612
ID AB862612 standard; protein; 346 AA.
XX
AC AB862612;
XX
DT 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 14628.
XX
DE Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
PD 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US009231.
XX
PF WPI; 2001-656860/75.
PR N-PSDB; ABL06715.
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions.
XX
XX Disclosure; SEQ ID NO 14628; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins (AB857737-
XX AB872072). The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 346 AA;
XX
XX Alignment Scores:
XX Pred. No.: 1,41e-24 Length: 346
XX Score: 359.50 Matches: 82
XX Percent Similarity: 46.1% Conservative: 35
```

```
Best Local Similarity: 32.3% Mismatches: 88
Query Match: 8.2% Indels: 49
DB: 4 Gaps: 5
US-10-768-158-1 (1-2419) x AB862612 (1-346)
Qy 84 TTCATGGCGTGGCG-----CTGCCGCC 107
Db 27 PheHisGlyGluArgThrGlyPheValGlnValGlySerGluGlyTyrPhePheProHis 46
Qy 108 TTCTCCCGCGGGAAGTGGAGAGATCGCCAACTTCCTCCGGTGGCGCCAGCGAGCTGTGG 167
Db 47 LysTyrIysAspGluAlaGluArgTyrTyrAsnPheGluAlaArgProAspAspValTyr 66
Qy 168 ATCGTCACTACCCCAAGTCCGCGACCGAGCTTGTCTGTCAGGAGGTGTCTACTTGGTGAGC 227
Db 67 IleAlaThrValProArgSerGlyThrTrpThrGlnLeuLeuIleTyrLeuValAla 86
Qy 228 CAGGCGCTGACCCCGATGAGATCGGCTTGATGAACATCCAGCAGGAGCTCCCGGTCCTG 287
Db 87 AsnGlyLeuAspPheGluHisAlaGlnGluArgProLeuThrGluArgPheProPhePhe 106
Qy 288 GAGTACCCA----- 296
Db 107 GluPheProLeuPheValHisProLysIleLysGluGluGlnGluGluAsnArgAsp 126
Qy 297 -----CAGCGGGCGCTGGACATCATCAAG 320
Db 127 SerAlaGluAlaLeuGluPheIleGluLysIleAlaArgProGlyTyrGluAlaLeuSer 146
Qy 321 GAACTG-----ACCTCTCCCGCTCATCAAGAGCCACCTGCCCTACCGCTTCTCTGCC 374
Db 147 GluIleProArgSerGlnArgPheIleLysThrHisPheProPheSerLeuMetPro 166
Qy 375 TCTGACCTCCACAATGGAGACTCCAAGTCTATATGCTGCGCAACCCCAAGGATCTG 434
Db 167 ProSerValLeuGluLysLysCysLysValIleTyrValValArgAspProLysAspVal 186
Qy 435 GTGGTGTCTTATTATCATGTTCCACCGCTCTCTGGGACCATGAGCTACCGAGGACCTTT 494
Db 187 AlaValSerTyrTyrHisLeuAsnArgLeuPheArgThrGlnGlyTyrValGlyAspPhe 206
Qy 495 CAAGAATTCTGCGGAGGTTTATGAATGATAAGCTGGCTACGCTCCTCG----- 545
Db 207 GluArgTyrTrpHisTyrPheGlnAsn-----GlyLeuAsnProTyrLeuProTyr 223
Qy 546 TTTGAGCAGCTGCGAGGAGTTCTGGGAGCACCGCATGGACTCGAACGTGCTTTTCTCAAG 605
Db 224 TyrSerHisValLysGluAlaArgGluHisAlaHisLeuSerAsnValLeuPheLeuArg 243
Qy 606 TATGAGACATGTCATCGGACCTGGTGTGACCATGATGTTGGAGCAGCTGGCCAGATTCTCTGGG 665
Db 244 TyrGluAspMetLeuAlaAspLeuProGlyAlaIleAsnSerIleAlaSerPheLeuGlu 263
Qy 666 GTGTCTCTGTGACAGGCCCGAGCTGGAAGCCCTGACGAGGAC 707
Db 264 CysProLysProGluAspMetAspArgLeuLeuAspHis 277
RESULT 100
AB862695
ID AB862695 standard; protein; 254 AA.
XX
XX AB862695;
XX
XX 18-NOV-2004 (first entry)
XX
XX Human diagnostic and therapeutic pprotein SEQ ID NO:2944.
XX
XX gene therapy; human diagnostic and therapeutic polynucleotide; dithp.
XX
XX Homo sapiens.
XX
XX WO2004023973-A2.
XX
XX
```

XX PD 25-MAR-2004.

XX PF 12-SEP-2003; 2003WO-US028227.

XX PR 12-SEP-2002; 2002US-0410259P.

XX PR 12-SEP-2002; 2002US-0410260P.

XX PA (INCY-) INCYTE CORP.

XX PI Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F; Harthorne TA, Suchorski MT, Altus CM, Pitts SJ, Elder LV; Money EM, Deleane AM, Panesar IS, Banville SC, Reddy TP; Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH; Peralta CA, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL; Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES; Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D; Patry S, Shi X, Suarez CJ;

XX WPI; 2004-329368/30.

XX N-PSDB; ACN41347.

XX PT New diagnostic and therapeutic polynucleotides and polypeptides, useful in diagnosing a condition, disease or disorder associated with human molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or in gene mapping.

XX PS Claim 27; Page; 190pp; English.

XX CC The invention relates to novel diagnostic and therapeutic polynucleotides selected from one of the 2722 sequences defined in the specification. A polynucleotide of the invention may have a use in gene therapy. The human diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be used to diagnose a particular condition, disease or disorder associated with human molecules, e.g. cell proliferative disorders, autoimmune/inflammatory disorder, developmental disorder, endocrine disorder, neurological disorders, gastrointestinal disorders, or infections caused by virus, bacteria, fungi or parasite. The dithp molecules may also be used in genetic mapping, in identifying individuals from minute biological samples, in detecting single nucleotide polymorphisms, as molecular weight markers, and for somatic or germline gene therapy. The present sequence represents a dithp protein of the invention. Note: The sequence data for this patent is not represented in the printed specification, but was obtained in electronic format directly from WIPO at www.wipo.int/pct/en/sequences/listing.htm

XX SQ Sequence 254 AA;

Alignment Scores:

Pred. No.:	1.86e-24	Length:	254
Score:	357.50	Matches:	81
Percent Similarity:	46.9%	Conservative:	46
Best Local Similarity:	29.9%	Mismatches:	93
Query Match:	8.1%	Indels:	51
DB:	8	Gaps:	4

US-10-768-158-1 (1-2419) x ABM82695 (1-254)

Qy 90 GCGTGGCGCTCGCCCTTCTGCGCGGGAAGATGGAGGAGATCGCAACTCCCGTG 149

Db 17 GlyValProLeuIleLysTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla 36

Qy 150 CGGCCACGACGTGTGGATCGTACCTACCCCAAGTCCCGCAGCTTCTGTCGAGGAG 209

Db 37 ArgProAspPleuLeuIleSerThrTyrProLysSerGlyThrTyrTrpValSerGln 56

Qy 210 GTGCTCTACTTGTGAGCGGGCGCTGACCCCGATGATGATGCGCTTGATGAACATCGAC 269

Db 57 IleLeuAspMetIleTyrGlnGlyGlyAsp----- 66

Qy 270 GAGCAGCTCCCGGCTCTGGAGTACCCACAGCGGGCGCTGGACATCATCAAGGAATGACC 329

Db 66 ----- 66

Qy 330 TCTCCCGCCTCATCAAGAGCCACCTGCCCTTCTGCGCTTCTGACCTCCACAAT 389

Db 67 -----LeuGluLysCysHisArg-----ValProGlnThrLeuLeuAsp 79

Qy 390 GGAGACTCCAGGTCATCTATATGGCTCGCAACCCCAAGGATCTGGTGGTCTTATAT 449

Db 80 GlnLysValLysValTyrValAlaArgAsnAlaLysAspValAlaValSerTyrTyr 99

Qy 450 CAGTTCCACCGCTCTCTCGCGGACCATGAGCTACCGAGCCACCTTTCAAGAAATTCTGCCGG 509

Db 100 HisPheTyrHisMetAlaLysValHisProGluProGlyThrTrpAspSerPheLeuGlu 119

Qy 510 AGTTTATGAATATAAGCTGGGTACGGTCTCTGGTTTGAGCAGCTGCAGAGTTCGG 569

Db 120 LysPheMetValGlyGluValSerTyrGlySerTrpTyrGlnHisValGlnGluTrpTrp 139

Qy 570 GAGCAGCGCTGGACTCGAACGCTGCTTTTCTCAAGTATGAAGACATGCATCGGGACCTG 629

Db 140 GluLeuSerArgThrHisProValLeuTyrLeuPheTyrGluAspMetLysGluAsnPro 159

Qy 630 GTGACGATGTGGAGCAGCTGGCCAGATTCTGGGGGTCTCTGTGACAAGGCCACAGCTG 689

Db 160 LysArgGluLeuGlnLysIleLeuGluPheValGlyArgSerLeuProGluGluThrVal 179

Qy 690 GAAGCCCTGACGGAGCAGCTGC-----CACCAGCTGGTGGACCAAGTGTCTGCAACGCT 740

Db 180 AspPheValValGlnHisThrSerPheLysGluMetLysLysAsnProMetThrAsnTyr 199

Qy 741 GAGCCCTGCCC-----GTGGCGCGGGAAGA 767

Db 200 ThrThrValProGlnGluPheMetAspHisSerIleSerProPheMetArgLysGlyMet 219

Qy 768 GTTGGGCTGTGGAGGACATCTTCACCGTCTCCATGAATGAGAAGTTTGACTTGGTGAT 827

Db 220 AlaGlyAspTrpLysThrPheThrValAlaGlnAsnGluArgPheAspAlaAspTyr 239

Qy 828 AAACAGAAAGATGGGAAAGTGTGACCTCACGTTT 860

Db 240 AlaGluLysMetAlaGlyCysSerLeuSerPhe 250

Search completed: May 17, 2006, 09:32:02

Job time : 450.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 17, 2006, 09:18:30 ; Search time 141.6 Seconds
(without alignments)
4740.705 Million cell updates/sec

Title: US-10-768-158-1

Perfect score: 4401

Sequence: 1 ggcagcgagcgagcgagcgagc.....aaaaaaaaaaaaaaaaaaaaa 2419

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2849598 seqs, 92501592 residues

Total number of hits satisfying chosen parameters: 5699196

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-Q=/abs/ABSSWEB spool/US10768158/runat_16052006_163422_5282/app_query.fasta_1
-DB=UniProt -QFMT=fastan -SUFFIX=n2p.rup -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40 cdi -LIST=100
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=100 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abes02p
-USER=US10768158 @CCN_1_1_466 @runat_16052006_163422_5282 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

UniProt_7.2.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1530	34.8	284	1	ST4A1_HUMAN
2	1525	34.7	284	2	Q5R4X3_PONPY
3	1508	34.3	284	2	Q85L7_RABIT
4	1506	34.2	284	1	ST4A1_MOUSE
5	1506	34.2	284	1	ST4A1_MOUSE
6	1506	34.2	284	2	Q3TXY5_MOUSE
7	1483	33.7	284	2	Q8BQV8_MOUSE
8	1385	31.5	284	2	Q6BQV3_XENLA
9	1382	31.4	284	2	Q52KW3_XENLA
10	1313	29.8	284	2	Q4SRF8_TETNG
11	1129.5	25.7	218	2	Q5RCAB_PONPY
12	495	11.2	298	2	Q9R1S5_MOUSE
13	489	11.1	291	1	ST1A1_MOUSE
14	480	10.9	296	1	ST1C1_MOUSE
15	476	10.8	296	2	Q3UNA9_MOUSE
16	472	10.7	295	1	ST1A3_HUMAN

17	467	10.6	296	1	STK1_RAT	Q9WUW8	rattus norv
18	467	10.6	296	2	Q3ZAV3_RAT	Q3ZAV3	rattus norv
19	466	10.6	296	2	Q642G8_RAT	Q642G8	rattus norv
20	465	10.6	296	2	Q569D0_RAT	Q569D0	rattus norv
21	464	10.5	295	2	Q2TAB3_HUMAN	Q2TAB3	homo sapien
22	464	10.5	296	1	ST1B1_CHICK	Q8jg30	gallus gall
23	463	10.5	298	2	Q58CV8_BOVIN	Q58CV8	bos taurus
24	458	10.4	291	1	ST1A1_RAT	PI7988	rattus norv
25	458	10.4	291	2	Q548D2_RAT	Q548D2	rattus norv
26	456	10.4	295	1	ST1A2_HUMAN	P50226	homo sapien
27	456	10.4	295	2	Q6Y0X5_PIG	Q6Y0X5	sus scrofa
28	456	10.4	296	1	STK2_RAT	Q9WUW9	rattus norv
29	453.5	10.3	296	2	Q95J76_TRIVU	Q95J76	trichosurus
30	451	10.2	295	2	ST1A1_HUMAN	P50225	homo sapien
31	451	10.2	295	1	Q2NL71_HUMAN	P02171	homo sapien
32	451	10.2	295	2	Q3ZC30_BOVIN	Q3ZC30	bos taurus
33	450	10.2	295	1	ST1A1_CANFA	Q29476	canis famil
34	449	10.2	295	2	Q95KM3_RABIT	Q95KM3	oryctolagus
35	449	10.2	296	1	ST1C1_HUMAN	O00338	homo sapien
36	449	10.2	302	1	ST1C2_HUMAN	O75997	homo sapien
37	447	10.2	295	2	Q95MF8_PIG	P52847	sus scrofa
38	446.5	10.1	299	1	ST1B1_RAT	Q95Mf8	rattus norv
39	446	10.1	295	1	ST1A1_MACFA	P52846	macaca fasc
40	446	10.1	295	1	ST1E1_BOVIN	P19217	bos taurus
41	443.5	10.1	296	1	ST1B1_HUMAN	Q43704	homo sapien
42	443	10.1	296	1	ST1B1_CANFA	Q95Jd5	canis famil
43	442.5	10.1	296	1	ST1C1_RABIT	Q46503	oryctolagus
44	442	10.0	292	2	Q9XT99_RABIT	Q9XT99	oryctolagus
45	440.5	10.0	297	2	Q6QP49_XENLA	Q6QP49	xenopus lae
46	436.5	9.9	296	2	Q3T0Y3_BOVIN	Q3T0Y3	bos taurus
47	435	9.9	295	1	ST1E2_RAT	P52845	rattus norv
48	435	9.9	295	1	ST1E6_RAT	P49890	rattus norv
49	434.5	9.9	299	1	ST1S1_BRARE	Q6ph37	brachydanio
50	434	9.9	294	1	ST1E1_HUMAN	P49888	homo sapien
51	434	9.9	294	2	Q53X91_HUMAN	Q53X91	homo sapien
52	432.5	9.8	301	2	Q6DHG7_BRARE	Q6dhg7	brachydanio
53	432	9.8	295	2	Q9D565_MOUSE	Q9D566	mus musculus
54	431.5	9.8	287	2	Q5HZV7_XENTR	Q5hsv7	xenopus tro
55	431	9.8	295	1	ST1E1_MOUSE	P49891	mus musculus
56	431	9.8	295	2	Q8JZX7_MOUSE	Q8jzx7	mus musculus
57	430	9.8	295	2	Q9SND5_RAT	Q9Snd5	rattus norv
58	429.5	9.8	299	1	ST1B1_MOUSE	Q9qwg7	mus musculus
59	429	9.7	263	2	Q91W13_MOUSE	Q91w19	mus musculus
60	429	9.7	295	1	ST1E3_RAT	P49889	rattus norv
61	429	9.7	295	2	Q3T0S9_BOVIN	Q3T0S9	bos taurus
62	428	9.7	307	2	Q90WR6_CHICK	Q90wr6	gallus gall
63	427.5	9.7	301	1	ST1S3_BRARE	Q7t2v2	brachydanio
64	427.5	9.7	301	1	ST2S2_BRARE	Q7Zus4	brachydanio
65	427.5	9.7	302	2	Q4FZP1_XENLA	Q4fzp1	xenopus lae
66	423	9.6	294	2	Q2XV97_MACFA	Q2xv97	macaca fasc
67	422	9.6	304	2	Q6IMI6_HUMAN	Q6imi6	homo sapien
68	421.5	9.6	304	2	Q6XZC1_BRARE	Q6xzcl	brachydanio
69	421	9.6	295	1	ST1E1_RAT	P52844	rattus norv
70	420	9.5	295	2	Q95J76_ORNAN	Q95j76	ornithorhyn
71	419.5	9.5	285	2	Q3HYK0_STRPU	Q3hyk0	strongyloce
72	419.5	9.5	304	1	SUAC_RAT	P50237	rattus norv
73	419.5	9.5	304	2	Q5M8B5_RAT	Q5m8b5	rattus norv
74	416.5	9.5	287	2	Q6P8G4_XENTR	Q6p8g4	xenopus tro
75	414.5	9.4	276	2	Q68EVA_XENLA	Q68ev4	xenopus lae
76	413	9.4	295	2	Q35401_MOUSE	Q35401	mus musculus
77	413	9.4	303	2	Q5EAWO_XENLA	Q5Eaw0	xenopus lae
78	412.5	9.4	304	2	Q80VR3_MOUSE	Q80vr3	mus musculus
79	412	9.4	294	1	ST1A1_BOVIN	P50227	bos taurus
80	412	9.4	294	2	Q7S293_XENLA	Q7sz93	xenopus lae
81	412	9.4	295	2	Q95JD6_CANFA	Q95jd6	canis famil
82	412	9.4	295	2	Q3UZ26_MOUSE	Q3uz26	mus musculus
83	412	9.4	295	2	Q6NZD1_MOUSE	Q6nzd1	mus musculus
84	412	9.4	308	2	Q9Z1G0_RAT	Q9z1g0	rattus norv
85	412	9.4	309	2	Q9R2C2_MOUSE	Q9r2c2	mus musculus
86	411	9.3	305	2	Q66KW4_XENLA	Q66kw4	xenopus lae
87	410.5	9.3	297	2	Q5H2U0_XENTR	Q5hzu0	xenopus tro
88	408.5	9.3	296	1	ST1E1_CANPO	P49887	cavia porce
89	408.5	9.3	304	2	O70262_MOUSE	O70262	mus musculus

90 405.5 9.2 309 2 Q4RG19 TETNG
 91 405 9.2 316 2 Q9VHHO DROME
 92 404.5 9.2 287 2 Q6DJL6 XENLA
 93 402 9.1 287 2 Q5M7C2 XENLA
 94 401.5 9.1 293 2 Q49IK6 BRARE
 95 401 9.1 304 2 Q6IMI5 HUMAN
 96 397.5 9.0 301 1 SF3A1_EABIT
 97 395.5 9.0 307 2 Q53SGA HUMAN
 98 394 8.9 304 2 Q6DJ68 XENR
 99 393.5 8.9 304 2 Q7ZYH0 XENLA
 100 392.5 8.9 299 2 Q6GL93 XENR

ALIGNMENTS

RESULT 1

ST4A1 HUMAN STANDARD; PRT; 284 AA.
 AC Q9BR01; O43728; integrated into UniProtKB/Swiss-Prot.
 DT 02-AUG-2002, sequence version 2.
 DT 07-FEB-2006, entry version 33.
 DE Sulfotransferase 4A1 (EC 2.8.2.-) (Brain sulfotransferase-like
 DE protein) (hBR-STL) (hBR-STL-1) (Nervous system sulfotransferase)
 DE (NST).
 GN Name=SLT4A1; Synonyms=SULTX3;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE (ISOFORM 1), AND CHARACTERIZATION.
 RC TISSUE=Brain;
 RX MEDLINE=20164074; PubMed=10698717; DOI=10.1042/0264-6021.3460857;
 RA Palany C.N., Xie X., Wang J., Ferrer J., Falany J.L.;
 RT "Molecular cloning and expression of novel sulphotransferase-like
 RT cDNAs from human and rat brain."
 RL Biochem. J. 346:857-864(2000).
 RN [2]
 RP NUCLEOTIDE SEQUENCE (ISOFORM 1), AND CHARACTERIZATION.
 RC TISSUE=Brain;
 RX MEDLINE=22035352; PubMed=12039030; DOI=10.1016/S0378-1119(02)00431-6;
 RA Sakakibara Y., Suiko M., Pai T.G., Nakayama T., Takami Y.,
 RA Katsufuchi J., Liu M.-C.;
 RT "Highly conserved mouse and human brain sulfotransferases: molecular
 RT cloning, expression, and functional characterization."
 RL Gene 285:35-47(2002).
 RN [3]
 RP NUCLEOTIDE SEQUENCE (ISOFORM 1).
 RC TISSUE=Brain;
 RA Martin S.C., Farb D.H.;
 RT "Molecular identification of a human nervous system cytoplasmic
 RT sulfotransferase, NST."
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP NUCLEOTIDE SEQUENCE (ISOFORM 1).
 RC TISSUE=Brain;
 RA Walther S.E., Raftogiannis R.B.;
 RT "Molecular and physical characterization of human SLT4A1,
 RT representing a novel cytosolic sulfotransferase family."
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP NUCLEOTIDE SEQUENCE (ISOFORM 2).
 RC TISSUE=Testis;
 RA Goward M.E., Huckle E.J.;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).
 RX PubMed=15461802; DOI=10.1186/gb-2004-5-10-r84;
 RA Collins J.E., Wright C.L., Edwards C.A., Davis M.P., Grinham J.A.,
 RA Cole C.G., Goward M.E., Aguado B., Mallia M., Mokrab Y., Huckle E.J.,
 RA Beare D.M., Dunham I.,

RT "A genome annotation-driven approach to cloning the human ORFeome.";
 RL Genome Biol. 5:RESEARCH84.1-RESEARCH84.11(2004).
 RN [7]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RX MEDLINE=20057165; PubMed=10591208; DOI=10.1038/990031;
 RA Dunham I., Hunt A.R., Collins J.E., Bruskiewicz R., Beare D.M.,
 RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
 RA Baguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
 RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,
 RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
 RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,
 RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,
 RA Dhani P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,
 RA Evans K.L., Fey J.M., Fleming K., French L., Garner A.,
 RA Gilbert J.G.R., Goward M.E., Grafham D.V., Griffiths M.N.D., Hall C.,
 RA Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S.,
 RA Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,
 RA Laird G.K., Langford C.F., Leverhwa M.A., Lloyd C., Lloyd D.M.,
 RA Martyn I.D., Maehregheh-Mohammadi M., Matthews L.H., Mccann O.T.,
 RA Mcclay J.D., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,
 RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,
 RA Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Ross M.T.,
 RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,
 RA Soderlund C., Spragon L., Steward C.A., Suleston J.E., Swann R.M.,
 RA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,
 RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,
 RA Wright C.B., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,
 RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,
 RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyama S.,
 RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,
 RA Dorman A., Fang F., Fu Y., Hua A., Kenton S., Lai H., Lao H.I.,
 RA Lewis J., Lewis S., Lin S.-P., Loh P., Malaj E., Nguyen T., Pan H.,
 RA Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,
 RA Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,
 RA Zhan M., Zhang G., Chisoe S., Murray J., Miller N., Minx P.,
 RA Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,
 RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,
 RA Hinds K., Kemp K., Latreille P., Layman D., Oersky P., Rohlfing T.,
 RA Scheet P., Walker C., Wamsley A., Wohlmann P., Pepin K., Nelson J.,
 RA Korf I., Bedell J.A., Hillier L.W., Mardis E., Waterston R.,
 RA Wilson R., Emanuel B.S., Shaikh T., Kurahashi H., Saitta S.,
 RA Budarf M.L., McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E.,
 RA Edelman L., Kim U.J., Shizuya H., Simon M.I., Dumanski J.P.,
 RA Peyrard M., Kedra D., Seroussi E., Fransson I., Tapia I., Bruder C.E.,
 RA O'Brien K.P., Wilkinson P., Bodenteich A., Hartman K., Hu X.,
 RA Khan A.S., Lane L., Tilahun Y., Wright H.;
 RT "The DNA sequence of human chromosome 22."
 RL Nature 402:489-495(1999).
 RN [8]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORMS 1 AND 2).
 RC TISSUE=Fetal brain, Hippocampus, and Hypothalamus;
 RX MEDLINE=2388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Atlachul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshikiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC !- FUNCTION: May catalyze the sulfate conjugation of many drugs,

CC xenobiotic compounds, hormones, and neurotransmitters. Displays
CC activity towards L-tryptophan, tyrosine, estrone, p-
CC nitrophenol, 2-naphthylamine, and 2-naphthol. May have a role in the
CC sulfation of drugs and neurotransmitters in the CNS.
CC SUBCELLULAR LOCATION: Cytoplasm (By similarity).
CC ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q9BR01-1; Sequences=Displayed;
CC Name=2;
CC IsoId=Q9BR01-2; Sequences=VSP 006304;
CC Note=No experimental confirmation available;
CC TISSUE SPECIFICITY: Highly expressed in the cerebral cortex and
CC frontal lobe, slightly less in the cerebellum, occipital and
CC temporal lobes, relatively low in the medulla and putamen, and
CC lowest in the spinal cord. No expression detected in the pancreas
CC (Ref.1). Highly expressed in fetal brain and occipital lobe,
CC slightly less in the whole brain, frontal lobe, hippocampus, and
CC lung, very low expression in cerebellum, medulla oblongata,
CC temporal lobe, testis, kidney and appendix (Ref.2).
CC SIMILARITY: Belongs to the sulfotransferase family.
CC
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC EMBL; AF188698; AAF61197.1; -; mRNA.
CC EMBL; AF115311; AAF21970.1; -; mRNA.
CC EMBL; AF176342; AAK64595.1; -; mRNA.
CC EMBL; AF251263; AAF98152.1; -; mRNA.
CC EMBL; AL590119; CAC34872.1; -; mRNA.
CC EMBL; CR456588; CAG30474.1; -; mRNA.
CC EMBL; Z97055; CAB09788.1; -; Genomic DNA.
CC EMBL; BC022459; AAH22459.1; -; mRNA.
CC EMBL; BC028171; AAH28171.1; -; mRNA.
CC EMBL; BC030665; AAH30665.1; -; mRNA.
CC HSSP; P50224; 1CJM.
CC Ensembl; ENSG00000130540; Homo sapiens.
CC HGNC; HGNC:14903; SULT4A1.
CC MIM; 608359; gene.
CC GO; GO:0005829; C:cytosol; NAS.
CC GO; GO:0008146; F:sulfotransferase activity; NAS.
CC InterPro; IPR00863; Sulfotransferase.
CC Pfam; PF00685; Sulfotransferase; 1.
CC ProDom; PD001218; Sulfotransferase; 1.
CC KX Alternative splicing; Lipid metabolism; Steroid metabolism;
KW Transferase.
FT CHAIN 1 284 Sulfotransferase 4A1.
FT FTID=PRO_0000085167.
FT NP_BIND 246 254 FAPS (Potential).
FT VARSPPLIC 248 284 GRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDPYL ->
FT AHCVPARKIFLSW (in isoform 2).
FT /FTID=VSP_006304.
FT FT CONFLICT 55 56 KS -> P (in Ref. 8; AAH30665).
FT FT CONFLICT 239 239 N -> S (in Ref. 8; AAH22459).
SQ SEQUENCE 284 AA; 33085 MW; A6EA6844B66C400B CRC64;

Alignment Scores:
Pred. No.: 1.17e-99 Length: 284
Score: 1530.00 Matches: 284
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 34.8% Indels: 0
DB: 1 Gaps: 0

US-10-768-158-1 (1-2419) x ST4A1_HUMAN (1-284)

QY 21 ATGGCGGAGAGCGAGCGCGAGACCCCGCGGGGAGTTTCGAGAGCAAGTACTTC 80
Db 1 MetaAlaGluSerGluAlaGluThrProSerThrProGluGluSerLysThrPhe 20

QY 81 GAGTTCATGGCTGGCGCTCCGCGCTTCTGCGCGGGAAGATGAGGAGATCCGCAAC 140
Db 21 GluPheHisGlyValArgLeuProPheCysArgGlyLysMetGluGluAlaAsn 40

QY 141 TTCCCGGTGCGCGCCAGCGAGCGTGTGGATGTCCTACCTACCCCAAGTCGGCACCAAGCTTG 200
Db 41 PheProValArgProSerAspValTrpIleValThrTyrProLysSerGlyThrSerLeu 60

QY 201 CTGCAGGAGGTGGTCTACTTGGTGAGCGCGCGCTGACCCCGATGAGTGGCTTGATG 260
Db 61 LeuGlnGluValValTyrLeuValSerGlnGlyAlaAspProAspGluIleLeuMet 80

QY 261 AACATCGACGAGCGCTCCCGCTCTGCGAGTACCCACAGCCGGGCTCGACATCATCAAG 320
Db 81 AsnIleAspGluGlnLeuProValLeuGluTyrProGlnProGlyLeuAspIleIleLys 100

QY 321 GAACGTACCTCTCCCGCTCATCAAGACCGCTGCGCTTCTGCGCTTCTGCGCTTCTGAC 380
Db 101 GluLeuThrSerProArgLeuIleLysSerHisLeuProTyrArgPheLeuProSerAsp 120

QY 381 CTCACAAATGAGACTCCAAAGTCAATCATATGCTGCTGCAACCCCAAGATCTGCTGGTG 440
Db 121 LeuHisAsnGlyAspSerLysValIleTyrMetAlaArgAsnProLysAspLeuValVal 140

QY 441 TCTTATTATCATGTTCCACCGCTCTCTGCGGACCATGAGCTACCGAGGACCTTTCAAGAA 500
Db 141 SerTyrTyrGlnPheHisArgSerLeuArgThrMetSerTyrArgGlyThrPheGlnGlu 160

QY 501 TTCTCGCGGAGGTTTATGAATGATAAGCTGGGCTACGGCTCTCTGGTTTGAGCACCTGCAG 560
Db 161 PheCysArgArgPheMetAsnAspLysLeuGlyTyrGlySerTrpPheGluHisValGln 180

QY 561 GAGTCTCGGAGCACCGCATGAGTCCGAACTGCTGCTTTCTCAAGTATGAGACATGCGAT 620
Db 181 GluPheTrpGluHisArgMetAspSerAsnValLeuPheLeuLysTyrGluAspMetHis 200

QY 621 CGGACCTGCTGACGATGGTGCGAGCAGCTGGCCAGATTCTCTGGGGGTGCTCTGTGACAAG 680
Db 201 ArgAspLeuValThrMetValGluGlnLeuAlaArgPheLeuGlyValSerCysAspLys 220

QY 681 GCCAGCTGGAAAGCCCTGACCGAGCATGTCACAGCTGGTGGACAGTGTGCAACGCT 740
Db 221 AlaGlnLeuGluAlaLeuThrGluHisCysHisGlnLeuValAspGlnCysCysAsnAla 240

QY 741 GAGCCCTGCGCGCGCGGGAAGATGGGCTGTGGAAGGACATCTTCCACCGCTCTCC 800
Db 241 GluAlaLeuProValGlyArgGlyValGlyLeuTrpLysAspIlePheThrValSer 260

QY 801 ATGAATGAGAAGTTTACTTGGTGTATATAAAGAGATGGGAAAGTGTGACCTCACGTTT 860
Db 261 MetAsnGluLysPheAspLeuValTyrLysGlnLysMetGlyLysCysAspLeuThrPhe 280

QY 861 GACTTTTATTATA 872
Db 281 AspPheTyrLeu 284

RESULT 2
QSR4X3_PONPY PRELIMINARY; PRT; 284 AA.
ID QSR4X3_PONPY AC QSR4X3;
DT 21-DEC-2004, integrated into UniProtKB/TrEMBL.
DT 21-DEC-2004, sequence version 1.
DT 07-FEB-2006, entry version 7.
DE Hypothetical protein DKFZp459L1114.
GN Name=DKFZp459L1114;
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Cortex;
RG The German cDNA Consortium;
RA Bloeker H., Boecher M., Brandt P., Mewes H.W., Weil B., Amid C.,
RA Osanger A., Fobo G., Han M., Wiemann S.;

```

RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL, CR861116; CAH93193.1; -; mRNA.
DR GO: 0008146; F:sulfotransferase activity; IEA.
DR InterPro: IPR000863; Sulfotransferase.
DR Pfam: PF00685; Sulfotransferase_1; 1.
DR ProDom: PD001218; Sulfotransferase; 1.
KW Hypothetical protein.
SQ SEQUENCE 284 AA; 33111 MW; 531F6C10A3294B00 CRC64;

Alignment Scores:
Pred. No.: 2,65e-99 Length: 284
Score: 1525.00 Matches: 283
Percent Similarity: 99.6% Conservativeness: 0
Best Local Similarity: 99.6% Mismatches: 1
Query Match: 34.7% Indels: 0
DB: Gaps: 0

US-10-768-158-1 (1-2419) x Q865L7_PONPY (1-284)
QY 21 ATGCGGAGAGCGAGCGGAGACCCCGGAGTTCGAGGCAAGTACTTC 80
DB 1 MetProGluSerGluAlaGluThrProSerThrProGlyGluPheGluSerIysTyrPhe 20
QY 81 GAGTTCATGCGCGTCCGCGTCCGCGTTCGCGGGAAGATCGAGGAGATCCCAAC 140
DB 21 GluPheHisGlyValArgLeuProPheCysArgGlyLysMetGluGluIleAlaAsn 40
QY 141 TTCCCGGTCCGCGCGGAGCGAGCTGTGGATCGTCACTACCCCAAGTCCGGCAGCGTTG 200
DB 41 PheProValArgProSerAspValTrpIleValThrTyrProGlySerGlyThrSerLeu 60
QY 201 CTGCGAGAGGTGGTCTACTTGGTGGAGCGCGGCGTGAACCGATGAGATCGGCTTGATG 260
DB 61 LeuGlnGluValValTyrLeuValSerGlnGlyAlaAspProAspGluIleGlyLeuMet 80
QY 261 AACATCGACGACAGCTCCGCTCTGTGGATACCCACAGCGCGGCTCGACATCATCAAG 320
DB 81 AsnIleAspGluGlnLeuProValLeuGluTyrProGlnProGlyLeuAspIleIleLys 100
QY 321 GAACGTACCTCCCGGCGGAGAGTTCGAGGAGTTCGAGGAGTTCGAGGAGTTC 380
DB 101 GluLeuThrSerProArgLeuIleLysSerHisLeuProTyrArgPheLeuProSerAsp 120
QY 381 TCCCAATGGAGACTCCAAAGTCTATATATGCTCGCAACCCCAAGGATCTGGTGGTG 440
DB 121 LeuHisAsnGlyAspSerIysValIleTyrMetAlaArgAsnProLysAspLeuVal 140
QY 441 TCTTATTATCAGTTCACCGCTCTCTCGGACCATGAGCTACCGAGGACCTTTCAAGAA 500
DB 141 SerTyrTyrGlnPheHisArgSerLeuArgThrMetSerTyrArgGlyThrPheGlnGlu 160
QY 501 TTCTCCCGAGGTTTATGATGATGAGCTGAGCTACGCTCGCTCGTGGTTCGAGCAGTGCAG 560
DB 161 PheCysArgPheMetAsnAspLysLeuGlyTyrGlySerTrpPheGluHisValGln 180
QY 561 GAGTTCCTGGAGCACCGGATGAGCTCGAAGCTGCTTTTCTCAAGTATGAGACATGAT 620
DB 181 GluPheTrpGluHisArgMetAspSerAsnValLeuPheLeuLysTyrGluAspMetHis 200
QY 621 CGGACCTGGTGCAGATGGTGGAGCAGTGGCCAGATTCCTGGGGGTGTCCTGTGACCAAG 680
DB 201 ArgAspLeuValThrMetValGluGlnLeuAlaArgPheLeuGlyValSerCysAspLys 220
QY 681 GCCAGCTGGAGCCCTGACGAGCAGCTCCACACGAGCTGGTGGACAGCTGCTGCAACGCT 740
DB 221 AlaGlnLeuGluAlaLeuThrGluHisCysHisGlnLeuValAspGlnCysCysAsnAla 240
QY 741 GAGGCGCTCCCGGCGGAGAGTTCGAGGAGTTCGAGGAGTTCGAGGAGTTC 800

```

```

Db 241 GluAlaLeuProValGlyArgGlyValGlyLeuTrpLysAspIlePheThrValSer 260
QY 801 ATGATGAGAGTTCGACTTGGTCTATAAACAAGATGGGAAGTGTGACCTCACGTTT 860
DB 261 MetAsnGluLysPheAspLeuValTyrLysGlnLysMetGlyLysCysAspLeuThrPhe 280
QY 861 GACTTTTATTATTA 872
DB 281 AspPheTyrLeu 284

RESULT 3
Q865L7_RABIT
ID Q865L7_RABIT PRELIMINARY; PRT; 284 AA.
AC Q865L7_
DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2003, sequence version 1.
DT 07-FEB-2006, entry version 11.
DE Brain sulfotransferase 4A1.
GN Name=Sult4a1;
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Lagomorpha; Leporidae;
OC Oryctolagus.
OX NCBI TaxID=9986;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA He D., Falany C.N.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AY196782; AA045181.1; -; mRNA.
DR HSP; P50224; 1CUM.
DR GO: 0008146; F:sulfotransferase activity; IEA.
DR GO: 0016740; P:transferase activity; IEA.
DR InterPro: IPR000863; Sulfotransferase.
DR Pfam: PF00685; Sulfotransferase_1; 1.
DR ProDom: PD001218; Sulfotransferase; 1.
KW Transferase.
SQ SEQUENCE 284 AA; 33107 MW; 782B689407A090DE CRC64;

Alignment Scores:
Pred. No.: 4.26e-98 Length: 284
Score: 1508.00 Matches: 279
Percent Similarity: 98.9% Conservativeness: 2
Best Local Similarity: 98.2% Mismatches: 3
Query Match: 34.3% Indels: 0
DB: Gaps: 0

US-10-768-158-1 (1-2419) x Q865L7_RABIT (1-284)
QY 21 ATGCGGAGAGCGAGCGGAGACCCCGGAGTTCGAGGAGTTCGAGGAGTTC 80
DB 1 MetAlaGluSerGluAlaGluThrProSerThrProGlyGluPheGluSerIysTyrPhe 20
QY 81 GAGTTCATGCGGTGCGGCTCCGCGTCCGCGTTCGCGGGAAGATCGAGGAGTTC 140
DB 21 GluPheHisGlyValArgLeuProPheCysArgGlyLysMetGluGluIleAlaAsn 40
QY 141 TTCCCGGTCCGCGCGGAGCGAGCTGTGGATCGTCACTACCCCAAGTCCGGCAGCGTTG 200
DB 41 PheProValArgProSerAspValTrpIleValThrTyrProLysSerGlyThrSerLeu 60
QY 201 CTGCGAGAGGTGGTCTACTTGGTGGAGCGCGGCGTGAACCGATGAGATCGGCTTGATG 260
DB 61 LeuGlnGluValValTyrLeuValSerGlnGlyAlaAspProAspGluIleGlyLeuMet 80
QY 261 AACATCGAGCAGCTCCGCTCTGTGGATACCCACAGCGCGGCTCGACATCATCAAG 320
DB 81 AsnIleAspGluGlnLeuProValLeuGluTyrProGlnProGlyLeuAspIleIleLys 100
QY 321 GAACGTACCTCCCGGCGGAGAGTTCGAGGAGTTCGAGGAGTTCGAGGAGTTC 380

```


CC hypothalamus. Not detected in pancreas, liver, lung, intestine,
 CC kidney, uterus, adrenal gland, thymus, spleen, epididymis,
 CC testicle, and heart.
 CC -!- SIMILARITY: Belongs to the sulfotransferase family.
 CC
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC
 CC -----
 DR EMBL; AF059257; AAC63999.1; -; mRNA.
 DR EMBL; AK003034; BAB2522.1; -; mRNA.
 DR EMBL; AK010293; BAB26829.1; -; mRNA.
 DR EMBL; BC051132; AAH51132.1; -; mRNA.
 DR EMBL; BC054757; AAH54757.1; -; mRNA.
 DR HSP; P50224; ICMJ.
 DR MGI; MGI:188971; Sult4a1.
 DR GO; GO:0005737; C:cytoplasm; TAS.
 DR GO; GO:0008146; F:sulfotransferase activity; IDA.
 DR GO; GO:0006790; P:sulfur metabolism; IDA.
 DR InterPro; IPR000863; Sulfotransferase.
 DR Pfam; PF00685; Sulfotransferase_1; 1.
 KW Alternative splicing; Lipid metabolism; Steroid metabolism;
 KW Transferase.
 FT CHAIN 1 284 Sulfotransferase 4A1.
 FT NP_BIND 246 254 PAPS (Potential).
 FT VARSP_LIC 248 284 GRVGLWKDIFVSNKEFDLVYKQWKGCDLTFDFYL ->
 FT AHCLFTKIALRWRCRGSGSLHCLDLVHVTA (in
 FT isoform 2).
 FT /FTID=VSP_006305.
 FT CONFLICT 6 6 A -> R (in Ref. 2; BAB2522).
 FT SEQUENCE 284 AA; 33054 MW; FCAE940F7219B6FC CRC64;
 Alignment Scores:
 Score: 5.91e-98 Length: 284
 Pred. No.: 1506.00 Matches: 278
 Percent Similarity: 99.3% Conservative: 4
 Best Local Similarity: 97.9% Mismatches: 2
 Query Match: 34.2% Indels: 0
 DB: 1 Gaps: 0
 US-10-768-158-1 (1-2419) x ST4A1_MOUSE (1-284)
 QY 21 ATGGCGGAGAGCGGCGGAGACCCAGCACCCTCCGCGGGAGTTTCGAGAGCAAGTACTTC 80
 DB 1 MetAlaGluSerGluAlaGluThrProGlyThrProGlyGluPheGluSerTyrPhe 20
 QY 81 GAGTTCCATGCGTCCGCTCCGCGCTTCTCCCGCGGGAAGATGGAGGAGATCGCCCAAC 140
 DB 21 GluPheHisGlyValArgLeuProProPheCysArgGlyLysMetGluAspIleAlaAsp 40
 QY 141 TTCGCGGTCCGCGGCGGAGCGTGTGGATCGTCACCTACCCCAAGTCCGCGCACCGCTTG 200
 DB 41 PheProValArgProSerAspValTrpIleValThrTyrProLysSerGlyThrSerLeu 60
 QY 201 CTCGAGGAGGTGGTCTACTTGGTGGACGAGCGGCGCTGACCCCGATGAGATCGGCTTGATG 260
 DB 61 LeuGlnGluValValTyrLeuValSerGlnGlyAlaAspProAspGluIleGlyLeuMet 80
 QY 261 AACATCGACGACGACTCCCGCTCTCGGAGTACCCACAGCCGCGGCTCGACATCATCAAG 320
 DB 81 AsnIleAspGluGlnLeuProValLeuGluTyrProGlnProGlyLeuAspIleIleLys 100
 QY 321 GAATGACCTCTCCCGCTCATCAAGAGCCACCTGCCCTACCGCTTCTGCCCTCTGAC 380
 DB 101 GluLeuThrSerProArgLeuIleLysSerHisLeuProTyrArgPheLeuProSerAsp 120
 QY 381 TCCCAATGAGAGCTCCAAGGTCTATATGCTGCTCGCAACCCCAAGATCTGGTGGTG 440
 DB 121 LeuHisAsnGlyAspSerLysValIleTyrMetAlaArgAsnProLysAspLeuValVal 140
 QY 441 TCTTATTATCAGTTCCACCGCTCTCTCGGACCATGAGCTACCGGAGCACCTTCAAGAA 500
 DB 141 SerTyrTyrGlnPheHisArgSerLeuArgThrMetSerTyrArgGlyThrPheGlnGlu 160

QY 501 TTCTGCGGAGGTTTATGAATGATAAGCTGGGCTACGGCTCCTGGTTTGGACAGTGCAG 560
 DB 161 PheCysArgArgPheMetAsnAspLysLeuGlyTyrGlySerTyrPheGluHisValGln 180
 QY 561 GAGTTCTGGGAGCACCGCATGAGCTGAACTGCTCTTTCTCAAGTATGAAGACATGCAT 620
 DB 181 GluPheTrpGluHisArgMetAspAlaAsnValLeuPheLeuLysTyrGluAspMetHis 200
 QY 621 CGGACCTGGTGACGATGGTGAGCAGCTGCGCCAGATTCCTGGGGGTGCTCTGTGACAAG 680
 DB 201 ArgAspLeuValThrMetValGlnLeuAlaArgPheLeuGlyValSerCysAspLys 220
 QY 681 GCCAGCTGGAGCCCTGACGAGCAGCTGCCACGAGCTGGTGGACAGTGTGCTGCAACGCT 740
 DB 221 AlaGlnLeuGluSerLeuIleGluHisCysHisGlnLeuValAspGlnCysCysAsnAla 240
 QY 741 GAGGCCCTGCCCGTGGCGGGAAGAGTGGGCTGTGGAGGACATCTTCACCGTCTCC 800
 DB 241 GluAlaLeuProValGlyArgGlyArgValGlyLeuTrpLysAspIlePheThrValSer 260
 QY 801 ATGAATGAGAAGTTTGAATGCTGTATAAACAGAGATGGGAAAGTGTGACCTCACGTTT 860
 DB 261 MetAsnGluLysPheAspLeuValTyrLysGlnLysMetGlyLysCysAspLeuThrPhe 280
 QY 861 GACTTTTATTTA 872
 DB 281 AspPheTyrLeu 284
 RESULT 5
 ID ST4A1_RAT STANDARD; PRT; 284 AA.
 AC P63047; O88872; Q91XS5; Q5CW7; Q9DC97;
 DT 31-AUG-2004, integrated into UniProtKB/Swiss-Prot.
 DT 31-AUG-2004, sequence version 1.
 DT 07-FEB-2006, entry version 14.
 DE Sulfotransferase 4A1 (EC 2.8.2.-) (Brain sulfotransferase-like
 DE protein) (rBR-STL) (Nervous system sulfotransferase) (NST).
 GN Name=Sult4a1; Synonyms=Sultx3;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [MRNA], TISSUE SPECIFICITY, DEVELOPMENTAL STAGE,
 RP AND SUBCELLULAR LOCATION.
 RC STRAIN=Sprague-Dawley; TISSUE=Brain;
 RX MEDLINE=20164074; PubMed=10698717; DOI=10.1042/0264-6021:3460857;
 RA Falany C.N., Xie X., Wang J., Ferrer J., Falany J.L.;
 RT "Molecular cloning and expression of novel sulphotransferase-like
 RT cDNAs from human and rat brain."
 RL Biochem. J. 346:857-864 (2000).
 RN [2]
 RP NUCLEOTIDE SEQUENCE [MRNA] OF 125-284.
 RC TISSUE=Brain;
 RA Martin S.C., Farb D.H.;
 RT "Molecular identification of a rat nervous system cytoplasmic
 RT sulfotransferase, NST."
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: May catalyze the sulfate conjugation of many drugs,
 CC xenobiotic compounds, hormones, and neurotransmitters. Displays
 CC activity towards L-triiodothyronine, thyroxine, estrone, p-
 CC nitrophenol, 2-naphthylamine, and 2-naphthol. May have a role in the
 CC sulfation of drugs and neurotransmitters in the CNS (By
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasm.
 CC -!- TISSUE SPECIFICITY: Expressed in the brain, not detected in the
 CC liver, kidney, spleen, heart, small intestine or testis.
 CC -!- DEVELOPMENTAL STAGE: Expressed at low levels in brains of 1-day
 CC old animals but increase to adult levels from 7-day old animals
 CC and remain at that level in adults.
 CC -!- SIMILARITY: Belongs to the sulfotransferase family.

```
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL: AF188699; AAF61198.1; -; mRNA.
DR EMBL: AF176343; AAK64596.1; -; mRNA.
DR HSP: F50224; ICDJ.
DR RGD: 69292; Sult4al.
DR InterPro: IPR000863; Sulfotransferase.
DR Pfam: PF00685; Sulfotransferase 1; 1.
DR Lipid metabolism; Steroid metabolism; Transferase.
KW CHAIN 1 284
FT FTID=PRO_0000085169.
SQ SEQUENCE 284 AA; 33054 MW; FCAE940F7219B6FC CRC64;

Alignment Scores:
Pred. No.: 5,91e-98 Length: 284
Score: 1506.00 Matches: 278
Percent Similarity: 99.3% Conservative: 4
Best Local Similarity: 97.9% Mismatches: 2
Query Match: 34.2% Indels: 0
DB: 1 Gaps: 0

US-10-768-158-1 (1-241) x ST4A1_RAT (1-284)
QY 21 ATGGCGGAGAGCGAGCGGCGGAGACCCCGGGGGAGTTCGAGGCAAGTACTTC 80
DB 1 MetAlaGluSerGluAlaGluThrProGlyThrProGlyGluPheGluSerIysThrPhe 20
QY 81 GAGTTCCATGGCGTGGCGTGGCGGCTTCTGCGCGGGAAGATGGAGGAGATCCCAAC 140
DB 21 GluPheHisGlyValArgLeuProPheCysArgGlyLysMetGluAspIleAlaAsp 40
QY 141 TTCCCGGTGGCGCCAGCGAGTGGATCGTACCTACCCCAAGTCGGGACGACGCTTG 200
DB 41 PheProValArgProSerAspValThrIleValThrIleValThrIleValThrSerLeu 60
QY 201 CTGACGAGGAGTGGTCTACTTGGTAGCGAGCGGCGTGCACCCCGATGAGATCGGCTTGATG 260
DB 61 LeuGlnGluValValThrLeuValSerGlnGlyAlaAspProAspGluIleGlyLeuMet 80
QY 261 AACATCGACGACGCTCCGGTCTCTGAGTACCCACACGCGGCTCGACATCATCAAG 320
DB 81 AsnIleAspGluGlnLeuProValLeuGluThrProGlnProGlyLeuAspIleIleLys 100
QY 321 GAATGACCTCTCCCGCTCATCAAGAGCACCTGCGCTACCGCTTCTGCCCTCTCAC 380
DB 101 GluLeuThrSerProArgLeuIleLysSerHisLeuProTyrArgPheLeuProSerAsp 120
QY 381 CTCCACAATGAGACTCCAAAGTCACTATATGCTCGCAACCCCAAGGATCTGGTGGTG 440
DB 121 LeuHisAsnGlyAspSerLysValIleThrMetAlaArgAsnProLysAspLeuValVal 140
QY 441 TCTTATTATCAGTTCACCGTCTCTCGGACCATGAGCTACCGAGGACCTTTCAGAA 500
DB 141 SerTyrTyrGlnPheHisArgSerLeuArgThrMetSerTyrArgGlyThrPheGlnGlu 160
QY 501 TTCTGCGGAGGTTTATCAATGATAGCTGGCTACCGCTCTGTTGAGCAGCTGAG 560
DB 161 PheCysArgArgPheMetAsnAspLysLeuGlyTyrGlySerTrpPheGluHisValGln 180
QY 561 GAGTTCCTGGGACACCGATGACATCGCAACGTCGCTTTTCTCAAGTATGAGACATGCAT 620
DB 181 GluPheTrpGluHisArgMetAspAlaAsnValLeuPheLeuLysTyrGluAspMetHis 200
QY 621 CGGACCTGGTACGATGGTGAGCAGTGGCCAGATTCTGGGGGCTCTCTGTGACAAAG 680
DB 201 ArgAspLeuValThrMetValGluGlnLeuAlaArgPheLeuGlyValSerCysAspLys 220
QY 681 GCCACGCTGGAGCCCTGACGCGGACGTCGCCACGCTGGTGGACGCTGCTGCAACCT 740
DB 221 AlaGlnLeuGluSerLeuIleGluHisCysHisGlnLeuValAspGlnCysCysAsnAla 240
```

```
QY 741 GAGGCCCTGCCCGTGGCGCGGGAAGAGTGGCGCTGTGGAGGACATCTTACCCGCTCC 800
DB 241 GluAlaLeuProValGlyArgGlyArgValGlyLeuThrLysAspIlePheThrValSer 260
QY 801 ATGAATGAGAAGTTGACTTGGTGTATATAACAGAGATGGGAAAGTGTGACCTCAGTTT 860
DB 261 MetAsnGluLysPheAspLeuValTyrLysGlnLysMetGlyLysCysAspLeuThrPhe 280
QY 861 GACTTTTATTTA 872
DB 281 AspPheTyrLeu 284

RESULT 6
Q3TXV5_MOUSE
ID Q3TXV5_MOUSE PRELIMINARY; PRT; 284 AA.
AC Q3TXV5
DT 11-OCT-2005, integrated into UniProtKB/TrEMBL.
DT 11-OCT-2005, sequence version 1.
DT 07-FEB-2006, entry version 5.
DE Visual cortex cDNA, RIKEN full-length enriched library,
DE clone:K530039L09 product:sulfotransferase family 4A, member 1, full
DE insert sequence.
DE Name=Sult4al;
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Visual cortex;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Methods Enzymol. 303:19-44 (1999).
RN [2]
NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Visual cortex;
RX PubMed=16141072; DOI=10.1126/science.11112014;
RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
RA Oyama K., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,
RA Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,
RA Davis M.J., Wilming L.G., Aidinis V., Allen J.E.,
RA Ambesi-Impombato A., Apweiler R., Aturaliya R.N., Bailey T.L.,
RA Bansal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
RA di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G.,
RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,
RA Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,
RA Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
RA Hill D., Humnietcki L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
RA Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,
RA Kitano H., Kollas G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
RA Matsuda H., Mattiuzzi S., Miki H., Mignone F., Miyake S., Morris K.,
RA Mortagui-Tabar S., Mulder N., Nakano N., Nakauchi H., Ng P.,
RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavoni G., Pesole G.,
RA Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
RA Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,
RA Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sesca L., Sheng Y.,
RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
RA Sperling S., Stupka E., Sugiyara K., Sultana R., Takenaka Y., Taki K.,
RA Tammoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,
RA Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,
RA Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
RA Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
```


QY 261 PACATCGAGCAGCTCCCGCTCTGGAGTACCCACAGCCGGGCTGGACATCATCAAG 320
Db |||||||
81 AsnIleAspGluInLeuProValLeuGluTyrProGlnProGlyLeuAspIleIleIys 100
QY 321 GAACTGACCTCTCCCGCTCATCAAGAGCCACCTACCGCTCTTCTGCGCTCTGAC 380
Db |||||||
101 GluLeuThrSerProArgLeuIleIysSerHisLeuProTyrArgPheLeuProSerAsp 120
QY 381 CTCCTAATGGAGATCCAAAGTCTATATGGTTCGCAACCCCAAGGATCTGGTGTG 440
Db |||||||
121 LeuHisAsnGlyAspSerLysValIleTyrMetAlaArgAsnProLysAspLeuVal 140
QY 441 TCTTATTATCAGTTCACCGCTCTCGCGGACCATGACTACCGGACCTTTCACGAA 500
Db |||||||
141 SerTyrTyrGlnPheHisArgSerLeuArgThrMetSerTyrArgGlyThrPheGlnGlu 160
QY 501 TTCTCCCGGAGGTTTATGAATGATTAAGCTGGGCTACGGCTCTGTTGTGAGCAGCTGCAG 560
Db |||||||
161 PheCysArgArgPheMetAsnAspLysIleuGlyTyrGlySerTrpPheGluHisValGln 180
QY 561 GAGTTCCTGGGACACCGCATGCGTCAAGCGTCTTTTCTCAAGTATGAAGACATGCAT 620
Db |||||||
181 GluPheTrpGluHisArgMetAspAlaAsnValLeuPheLeuLysTyrGluAspMethis 200
QY 621 CGGACCTGCTGCACATGGTGAGCAGCTGGCCAGATTCCTGGGGGTGCTCTGTGACAAAG 680
Db |||||||
201 ArgAspLeuValThrMetValGluGlnLeuAlaArgPheLeuGlyValSerCysAspIys 220
QY 681 GCCCAGCTGGAAGCCCTGACGAGCAGCTGCCACACGAGCTGGTGACAGTGTGCAACGCT 740
Db |||||||
221 AlaGlnLeuGluSerLeuIleGluHisCysHisGlnLeuValAspGlnCysAsnAla 240
QY 741 GAGCGCTGCGCGTGGCGGGGGAAGAGTGGGCTGTGGAGGACATCTTCCCGTCTCC 800
Db |||||||
241 GluAlaLeuProValGlyArgGlyArgValGlyLeuTrpLysAspIlePheThrValSer 260
QY 801 ATGATGAGAGTTTGACTGCTGGTGTATTAACAGAGATGGGAAAGTGTGACCTCACGTTT 860
Db |||||||
261 MetAsnGluLysPheAspLeuValTyrLysGlnLysMetGlyLysCysAspLeuThrPhe 280
QY 861 GACTTTTATTTA 872
Db |||||||
281 AspPheTyrLeu 284
RESULT 7
Q8BQV8 MOUSE PRELIMINARY; PRT; 284 AA.
AC Q8BQV8_1
DT 01-MAR-2003, integrated into UniProtKB/TrEMBL.
DT 01-MAR-2003, sequence version 1.
DT 07-FEB-2006, entry version 20.
DE Adult male corpora quadrigemina cDNA, RIKEN full-length enriched
DE library, clone:B203076P15 product:sulfotransferase family 4A, member
DE 1, full, insert sequence.
GN Name=Sult4a1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Corpora quadrigemina; STRAIN=C57BL/6J;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Methods Enzymol. 303:19-44(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
RX PubMed=16141072; DOI=10.1126/science.1112014;
RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,

RA Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,
RA Davis M.J., Wilming L.G., Aidinis V., Allen J.E.,
RA Ambesi-Impombato A., Apweiler R., Aturaliya R.N., Bailey T.L.,
RA Bansal K., Batsali L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
RA di Bernardo D., Down T., Engstrom P., Fagiolini M., Faulkner G.,
RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,
RA Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,
RA Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
RA Hill D., Huminick L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
RA Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,
RA Kitano H., Kollias G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
RA Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
RA Mottaqui-Fabriz S., Mulder N., Nakano N., Nakaguchi H., Ng P.,
RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavoni G., Pesole G.,
RA Petrovski N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
RA Roat B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,
RA Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sheng Y.,
RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
RA Sperling S., Stupka E., Sugiuira K., Sultana R., Takenaka Y., Taki K.,
RA Tammoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.B., Yagi K.,
RA Yananishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,
RA Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
RA Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
RA Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Ninomiya N.,
RA Nishio T., Okada M., Plessy C., Shibata K., Shiraki T., Suzuki S.,
RA Tagami M., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J.,
RA Hayashizaki Y.;
RT "The transcriptional landscape of the mammalian genome.";
RL Science 309:1559-1563(2005).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
RX PubMed=16141073; DOI=10.1126/science.1112009;
RG RIKEN Genome Exploration Research Group, and Genome Science Group
RG (Genome Network Core team) and the FANTOM Consortium;
RT "Antisense transcription in the Mammalian Transcriptome.";
RL Science 309:1564-1566(2005).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaide I., Osato N., Saito R., Suzuki H., Yananaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Bruscia V., Chochia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai T., Lenhard B., Shibata K., Shinagawa A.,

RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
[5]
RL NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gliss C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staabli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Guatrecich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
[6]
RL NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
RX MEDLINE=2499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayase N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
[7]
RL NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
[8]
RL NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kaga I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh K., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akaiura S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AK046368; BAC32692.1; -; mRNA.
DR HSSP; P50224; 1CJM.

DR Ensembl; ENSMUSG00000018865; Mus musculus.
DR MGI; MGI:1888971; Sult4a1.
DR GO; GO:0005737; Cytoplasm; TAS.
DR GO; GO:0008146; F:sulfotransferase activity; IDA.
DR GO; GO:0008146; F:sulfotransferase activity; RCA.
DR GO; GO:0006790; P:sulfur metabolism; IDA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Transferase.
SQ SEQUENCE 284 AA; 32972 MW; 38D7B556F2B9391E CRC64;
Alignment Scores:
Pred. No.: 2,53e-96 Length: 284
Score: 1483.00 Matches: 274
Percent Similarity: 97.9% Conservative: 4
Best Local Similarity: 96.5% Mismatches: 6
Query Match: 33.7% Indels: 0
DB: 2 Gaps: 0
US-10-768-158-1 (1-2419) x Q8BQV8_MOUSE (1-284)
QY 21 ATGCGGAGAGCGAGCGGAGACCCCGAGACCCCGGAGGAGTTCGAGAGCAAGTACTTTC 80
DB 1 MetGlyGluSerGluAlaGluThrProGlyThrProGlyGluPheGluSerLysTyrPhe 20
QY 81 GAGTTCATGCGCGTCCCGCTCTGCGCGGAGAGATGGAGGATCGCCCAAC 140
DB 21 GluPheHisGlyValArgValProPheCysArgGlyLysMetGluAspIleAlaGlu 40
QY 141 TTCCCGGTGCGCGCCGAGCGAGCTGTGGATCTACCTACCTACCCCAAGTCGGGACCAAGTTG 200
DB 41 PheProGlyArgProSerAspValTrpIleValThrTyrProLysSerGlyThrSerLeu 60
QY 201 CTGACGAGGTGGTCTACTTGGTGAGCAGCGGCGCTGACCCCGATGAGATCGGCTTGATG 260
DB 61 LeuGlnGluValValTyrLeuValSerGlnGlyAlaAspProAspGluIleGlyLeuMet 80
QY 261 AACATCGACGAGCGTCCCGCTCTGAGTACCCAGCCAGCGCGGCTGGACATCATCAAG 320
DB 81 AsnIleAspGluGlnLeuAlaValLeuGluTyrProGlnProGlyLeuAspIleLeuLys 100
QY 321 GAATGACCTCTCCCGCTCATCAAGAGCCAGCTCCCTACCGCTTCTCCCTCTGAC 380
DB 101 GluLeuThrSerProArgLeuIleLysSerHisLeuProTyrArgPheLeuProSerAsp 120
QY 391 CTCACAAATGAGACTCCAAAGTCACTATATGCTGCGAACCCCAAGGATCTGGTGTG 440
DB 121 LeuHisAsnGlyAspSerLysValIleTyrMetAlaArgAsnProLysAspLeuValVal 140
QY 441 TCTTATTATCAGTTCACCGCTCTCTGCGGACCATAGCTACCGAGGACCTTTCAAGAA 500
DB 141 SerTyrTyrGlnPheHisArgSerLeuArgThrMetSerTyrArgGlyThrPheGlnGlu 160
QY 501 TTCTGCGGAGGTTTATGAATGAATGAGTGGGCTACCGCTCTCGTGTGGACAGCTGCGAG 560
DB 161 PheCysArgArgPheMetAsnAspLysLeuGlyTyrGlySerTrpPheGluHisValGln 180
QY 561 GAGTTCGGGAGCGCGCATCGAGCTGACAGCTGCTTTTCTCAAGATGAGAGACATGAT 620
DB 181 GluPheTrpGluHisArgMetAspAlaAsnValLeuPheLeuLysTyrGluAspMethis 200
QY 621 CGGACCTGGTGAGCGTGGTGGAGCAGCTGCCAGATTCCTGGGGGTCTCTGTGACAAAG 680
DB 201 ArgAspLeuValThrMetValGluGlnLeuAlaArgPheLeuGlyValSerCysAspLys 220
QY 681 GCCCAGCTGGAAGCCCTGACCGAGCACTGCCACAGCTGGTGGAGCCAGTCTGCTGCAACGCT 740
DB 221 AlaGlnLeuGluSerLeuIleGluHisCysHisGlnLeuValAspGlnCysCysAsnAla 240
QY 741 GAGGCCCTGCGCGTGGGCGCGGAGAGAGTGTGGGTGTGGAAGGACATCTTACCGCTCC 800
DB 241 GluAlaLeuProValGlyArgGlyArgValGlyLeuTyrLysAspIlePheThrValSer 260

```
QY 801 ATGAATGAGAACTTTGACTTGGTGTATATAACAGAGATGGGAAAGTGTGACCTCAGCTTT 860
|||||
Db 261 MetAsnGluLysPheAspLeuValTyrLysGlnLysMetGlyLysCysAspLeuThrPhe 280
|||||
QY 861 GACTTTTATTATTA 872
|||||
Db 281 AspPheTyrLeu 284

RESULT 8
ID Q68EV3_XENLA PRELIMINARY; PRT; 284 AA.
AC Q68EV3;
DT 11-OCT-2004, integrated into UniProtKB/TrEMBL.
DT 11-OCT-2004, sequence version 1.
DT 07-FEB-2006, entry version 12.
DE MGC84327 protein.
GN Name=MGC84327;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Eye;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391 (2002).
[2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Eye;
RX MEDLINE=22389257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Frange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Eye;
RA Klein S., Gerhard D.S.; to the EMBL/GenBank/DBJ databases.
RL Submitted (AUG-2004)
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
DR EMBL; BC080097; AAH80097.1; -; mRNA.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
SQ SEQUENCE 284 AA; 33211 MW; 9344EE54A62729D3 CRC64;

Alignment Scores:
```

```
Pred. No.: 2.28e-89 Length: 284
Score: 1385.00 Matches: 253
Percent Similarity: 95.4% Conservative: 18
Best Local Similarity: 89.1% Mismatches: 13
Query Match: 31.5% Indels: 0
DB: 2 Gaps: 0

US-10-768-158-1 (1-2419) x O68EV3_XENLA (1-284)

QY 21 ATGCGGAGAGCGAGCGCGAGACCCAGCACCCCGGGGAGTTTCGAGAGCAAGTACTTC 80
Db 1 MetAlaGluSerGluAlaGluThrProSerThrProCysGluPheGluSerLysTyrPhe 20
QY 81 GAGTTCCATGCGTGGCGCTCCGCCCTTTCGCCGGGAAAGATGGAGAGATCCCCAC 140
|||||
Db 21 GluTyrAsnGlyIleArgLeuProPheCysArgGlyLysMetGluGluValSerAsp 40
QY 141 TTCCCGGTGGCGCCAGCGAGCTGTGGATCTCCTACCTACCCCAAGTCGCGCACCAAGCTG 200
|||||
Db 41 PheProValArgGluAsnAspIleTrpIleValThrTyrProLysSerGlyThrSerLeu 60
QY 201 CTGCAGGAGGTGGTCTACTTGTGAGCCAGGCGCTGACCCCGATGAGATCGGCTTGATG 260
Db 61 LeuGlnGluValValTyrLeuValSerGlnGlyAlaAspProAspGluIleGlyLeuMet 80
QY 261 AACATGACGAGCAGCTCCCGTCTGTGAGTACCCACAGCCGCGCTCGACATCATCAAG 320
Db 81 AsnIleAspGluGlnLeuProValLeuGluTyrProGlnProGlyLeuAspIleLys 100
QY 321 GAATGTACCTCTCCCGCTCATCAAGAGCCACCTGCCCTACCGCTTTCGCCCTCTGAC 380
Db 101 GluLeuThrSerProArgLeuIleLysSerHisLeuProTyrArgPheLeuProSerAsp 120
QY 381 CTCACAAATGAGACTCCAAAGTCTATATGCTCTCAACCCCAAGATCTGTGGTG 440
Db 121 LeuHisAsnGlyAsnSerLysValIleTyrMetAlaArgAsnProLysAspLeuVal 140
QY 441 TCTTATTATCAGTTCCACCGCTCTCTGCGGACCATAGCTACCGAGGACCTTTCACGAA 500
Db 141 SerTyrTyrGlnPheHisArgSerLeuArgThrMetSerTyrArgGlyThrPheGlnGlu 160
QY 501 TTCTGCGGAGGTTTATGAATGATGATGCTGGCTACCGCTCCTGGTTTGACACGTCGAG 560
Db 161 PheCysArgArgPheMetAsnAspLysLeuGlyTyrGlySerTyrPheAspHisValGln 180
QY 561 GAGTTCTGGGAGCACCGCATCGAATCGCTGCTTTTCTCAAGTATGAAGACATGCAT 620
Db 181 GluPheTyrAspHisArgLeuAspSerAsnValLeuPheLeuLysTyrGluAspMethis 200
QY 621 CGGACCTGTGTGACGATGGTGGAGCAGCTGCCAGATTCTCTGGGGGTCTCTGTGACAA 680
|||||
Db 201 LysAspLeuGlyThrMetValGlnLeuValArgPheLeuGlyValSerTyrAspLys 220
QY 681 GCCCAGCTGGGAAGCCCTGACCGAGCAGCTGCCACAGCTGGTGGACCACTGCTGCAACGCT 740
|||||
Db 221 AlaGlnLeuGluSerThrIleGluHisCysHisLeuLeuIleAspHisCysCysAsnAla 240
QY 741 GAGGCCCTGCGCGTGGCGCGGGAAGAGTGGGTGTGGAAGGACATCTTCAACCTCTCC 800
Db 241 GluAlaLeuProLleGlyArgGlyValGlyLeuTyrLysAspIlePheThrValSer 260
QY 801 ATGAATGAGAGTTTGACTTGGTGTATATAACAGAGATGGGAAAGTGTGACCTCAGCTTT 860
|||||
Db 261 MetAsnGluLysPheAspGlnValTyrLysGlnArgMetGlyLysLeuAspLeuAlaPhe 280
QY 861 GACTTTTATTATTA 872
|||||
Db 281 GluPheAsnLeu 284

RESULT 9
Q52KW3_XENLA PRELIMINARY; PRT; 284 AA.
ID Q52KW3_XENLA
AC Q52KW3;
```

DT 24-MAY-2005, integrated into UniProtKB/TrEMBL.
 DT 24-MAY-2005, sequence version 1.
 DT 07-FEB-2006, entry version 4.
 DE Hypothetical protein.
 OS *Xenopus laevis* (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
 OC Xenopodidae; Xenopus; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Eye;
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.,
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 RT initiative."
 RL Dev. Dyn. 225:384-391 (2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Eye;
 RX MEDLINE=2238257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.F., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Rubin A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.C., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalios D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.W., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Eye;
 RA Klein S., Gerhard D.S.;
 RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC -----
 DR EMBL; BC094166; AAH94166.1; -; mRNA.
 DR GO; GO:0008146; F:sulfotransferase activity; IEA.
 DR InterPro; IPR000863; Sulfotransferase.
 DR Pfam; PF00685; Sulfotransferase 1; 1.
 DR ProDom; PD001218; Sulfotransferase; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 284 AA; 33242 MW; FFA81B7E9A8BEAA9 CRC64;
 Alignment Scores:
 Pred. No.: 3,73e-89 Length: 284
 Score: 1382.00 Matches: 253
 Percent Similarity: 89.4% Conservatives: 18
 Best Local Similarity: 95.1% Mismatches: 13
 Query Match: 31.4% Indels: 0
 DB: 2 Gaps: 0
 US-10-768-158-1 (1-2419) x Q52KM3_XENLA (1-284)
 QY 21 ATGGCGGAGAGCGAGCCGAGACCCCGACCCCGGGGAGTTCGAGAGCAAGTACTTC 80
 DB 1 MetAlaGluSerGluAlaGluThrProSerThrProCysGluPheGluSerLysTyrPhe 20

```

RX PubMed=15496914; DOI=10.1038/nature03025;
RA Jallion O., Aury J.-M., Brunet F., Petit J.-L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicault S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Bienmont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Criau C., Duprat S., Brottier P., Coutanceau J.-P., Gouzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Wolff J.-N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crollius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RL the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope, Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; CAAS01014526; CAF96774.1; -; Genomic DNA.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransfer_1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
SQ SEQUENCE 284 AA; 32894 MW; 414AB27E7D8DF9A6 CRC64;

Alignment Scores:
Pred. No.: 2,94e-84 Length: 284
Score: 1313.00 Matches: 243
Percent Similarity: 93.0% Conservative: 21
Best Local Similarity: 85.6% Mismatches: 20
Query Match: 29.8% Indels: 0
DB: 2 Gaps: 0

US-10-768-158-1 (1-2419) x Q4SRF8_TETNG (1-284)
QY 21 ATGGCGAGAGCGCGGAGACCCAGACCCCGGGGAGTTCCGAGGCAAGTACTTC 80
DB 1 MetalaalaargGluuAaAspThrProserThrProleGluPheGluSerIysTyrPhe 20
QY 81 GAGTTCATGGCGTGGCGGCTGCGGCTTCTGCGCGGGAAGATGGAGGAGATGCCAAC 140
DB 21 GluPheAspGlyValargLeuProPhePheCysArgGlyLysMetGluGluileAlaGly 40
QY 141 TTCCCGGTGCGGCGGAGCGTGGATCGTCACCTACCCAGTCCGCGGACGAGCTTG 200
DB 41 PheSerLeuArgSerAspIlePrileValThrTyrProLysSerGlyThrSerLeu 60
QY 201 CTGACGAGAGTGGTCTACTTGGTAGCCAGCGGCGTGCACCCGATGAGATCGCTTGATG 260
DB 61 LeuGlnGluileValTyrLeuValSerGlnGlyAlaAspProAspGluileGlyLeuMet 80
QY 261 AACATCGACGACGCTCCCGGCTCTGAGTACCCACACCGCGGCGCTCGACATCATCAAG 320
DB 81 AsnileAspGluGlnLeuProValLeuGluTyrProGlnProGlyLeuAspIleileGln 100
QY 321 GAATGACCTCTCCCGGCTCATCAGAGCCACCTGCGGCTTCTGCGCTCTCTGAC 380
DB 101 GluLeuThrSerProArgLeuileLysSerHisLeuProTyrGlnPheLeuProThrAla 120
QY 381 CTCACAAATGAGAGTCCCAAGTCACTATATGGCTCGCAACCCCAAGATCTGGTGGTG 440
DB 121 LeuHisAsnGlyGluuAlaLysValIleTyrMetAlaArgAsnProLysAspLeuValVal 140
QY 441 TCTTATTATAGTTCCACCGCTCTCTGCGGACCATAGAGTACCGGAGCACCTTTCAAGAA 500
```

```

Db 141 SerTyrTyrGlnPheHisArgSerLeuArgThrMetSerTyrArgGlyThrPheGlnGlu 160
QY 501 TTCTCCCGGAGGTTTATGAATGATAAGCTGGGCTACGGCTCCTGGTTTGGAGCAGCTGCAG 560
DB 161 PheCysArgArgPheMetAsnAspLysLeuGlyTyrGlySerTyrPheGluHisValGln 180
QY 561 GAGTCTCGGAGCACCAGCATCGAATCGAATCGTCTTTTCTCAAGTATGAAGACATGCAT 620
DB 181 GluPheTyrGluHisArgMetAsnSerValLeuPheLeuLysTyrGluAspMetTyr 200
QY 621 CGGAGCCTGGTGCAGATGGTGCAGAGCTGGCCAGATTCCTGGGGGTCTCTGTGACAAAG 680
DB 201 LysAspLeuGlyThrPheValGluGluuAlaArgPheLeuGlyValSerCysAspLys 220
QY 681 GCCAGCTGGAAAGCCCTGACGCGACCTGCACACAGCTGGTGGACGACCTGCACACCGT 740
DB 221 AlaGlnLeuGluGlyLeuValGluSerCysAsnGlnLeuileGluGlnCysSerAsnSer 240
QY 741 GAGCCCTGCGGCGGCGGAGAGTGGGCTGTGGAGAGACATCTTCACCGTCTCC 800
DB 241 GluAlaLeuSerValCysArgGlyArgValGlyLeuTyrLysAspValPheThrValSer 260
QY 801 ATGAATGAGAAGTTGACTTGTCTGTATAAACAGAGATGGGAAAGTGTGACCTCAGCTT 860
DB 261 MetAsnAspLysPheAspThrIleTyrArgGlnLysMetCysLysSerAspLeuThrPhe 280
QY 861 GACTTTTATTATA 872
DB 281 AspPheGlyLeu 284

RESULT 11
Q5RCA8_PONPY PRELIMINARY; PRT; 218 AA.
AC Q5RCA8
DT 21-DEC-2004, integrated into UniProtKB/TrEMBL.
DT 21-DEC-2004, sequence version 1.
DE Hypothetical protein DKFP459I0535.
GN Name=DKFP459I0535;
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC Tissue=Cortex;
RG The German cDNA Consortium;
RA Koehrer K., Beyer A., Mewes H.W., Weil B., Amid C., Osanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; CR858370; CAH90599.1; -; mRNA.
DR SMR; Q5RCA8; 35-218.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransfer_1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Hypothetical protein.
SQ SEQUENCE 218 AA; 25328 MW; 87B9496A82B7B878 CRC64;

Alignment Scores:
Pred. No.: 2,86e-71 Length: 218
Score: 1129.50 Matches: 215
Percent Similarity: 81.7% Conservative: 0
Best Local Similarity: 81.7% Mismatches: 3
Query Match: 25.7% Indels: 45
DB: 2 Gaps: 1
```

```
US-10-768-158-1 (1-2419) x Q5RCN8_PONPY (1-218)
QY 88 ATGGCGTGCCTGCGCGCTTCTGCGCGGGAAGATGGAGAGATCGCCAACTTCCGG 147
DB 1 MetalaCyseGlyCysArgProSerAlaAlaGlyArgTrpArgSerProThrSerArg 20
QY 148 TGGCGCCAGGAGGTGTGGATCGTCACCTACCCCAAGTCGGGACACAGCTTGTGAGG 207
DB 21 CysGlyProAlaThrCysGlyLeuSerProThrProSerPro----- 34
QY 208 AGGTGGTCTACTTGTGTGAGCGGCGCTGACCCCGATGAGATCGGCTTGATGAACATCG 267
DB 34 ----- 34
QY 268 ACGAGCAGCTCCCGTCTCGAGTACCCACAGCCGGGCTGGACATCATCAAGAACTGA 327
DB 35 -----Glu-LeuT 37
QY 328 CCTCTCCCGCTCATCAAGACCCCTGCGCAACCCCAAGGATCTGGTGTCTATT 447
DB 57 snGlyAspSerLysValIleTyrMetAlaArgAsnProLysAspLeuValSerTyrT 77
QY 448 ATCAGTTCACCGCTCTCGCGACCATGAGTACCGAGGACCTTTCAGAAATCTGCC 507
DB 77 yrGlnPheHisArgSerLeuArgThrMetSerTyrArgGlyThrPheGlnGluPheCysA 97
QY 508 GGAGTTTATGATGATGAGTGGGCTACGGCTCTGCTTTGAGCAGCTGCGAGGTCT 567
DB 97 rGArgPheMetAsnAspLysLeuGlyTyrGlySerTrpPheGluHisValGlnGluPheT 117
QY 568 GGGACACCGGATGAGTACGACGTGCTTTTCTCAAGTATGAGACATGATCGGACC 627
DB 117 rpGluHisArgMetAspSerAsnValLeuPheLeuLysTyrGluAspMetHisArgAspL 137
QY 628 TGGTACGATGGTGGAGCAGCTGGCCAGATTCTGGGGGTGTCTGTGACAAAGCCGACC 687
DB 137 euValThrMetValGlnGlnLeuAlaArgPheLeuGlyValSerCysAspLysAlaGlnL 157
QY 688 TGGAGCCCTACGAGGACCTGCCACGAGCTGGTGGAGCAGTGTGCAACCTGAGGCC 747
DB 157 euGluAlaLeuThrGluHisCysHisGlnLeuValAspGlnCysCysAsnAlaGluAlaL 177
QY 748 TGCCCGTGGCGGGAAGAGTTGGCTGTGGAAGGACATCTTACCCGCTCTCCATGAATG 807
DB 177 euProValGlyArgGlyArgValGlyLeuTyrLysAspIlePheThrValSerMetAsnG 197
QY 808 AGAAGTTTGACTGGTGTATAACAGAGATGGGAAAGTGTGACCTCAGTTTGACTTTT 867
DB 197 luLysPheAspLeuValTyrLysGlnLysMetGlyLysCysAspLeuThrPheAspPheT 217
QY 868 ATTTA 872
DB 217 yrLeu 218
RESULT 12
Q9R1S5_MOUSE PRELIMINARY; PRT; 298 AA.
AC Q9R1S5;
DT 01-MAY-2000, integrated into UniProtKB/TrEMBL.
DT 01-MAY-2000, sequence version 1.
DT 07-FEB-2006, entry version 18.
DE Arylsulfotransferase STIA4.
GN Name=Sultial;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
```

```
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=129SV;
RA Honma W., Nagata K., Yamazoe Y.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; AB029487; BAA82321.1; -; Genomic_DNA.
CC PIR; S28183; S28183.
CC HSSP; P50224; 1CJM.
CC SMR; Q9R1S5; 12-298.
CC Ensembl; ENSMUSG00000030711; Mus musculus.
CC MGI; MGI:102896; Sultial.
CC GO; GO:0008146; F:sulfotransferase activity; IEA.
CC GO; GO:0016740; P:transferase activity; IEA.
CC InterPro; IPR000863; Sulfotransferase.
CC Pfam; PF00685; Sulfotransfer_1; 1.
CC ProDom; PD001218; Sulfotransferase; 1.
KW Transferase.
SQ SEQUENCE 298 AA; 34717 MW; E5100E784A263668 CRC64;
Alignment Scores:
Pred. No.: 334e-26 Length: 298
Score: 495.00 Matches: 104
Percent Similarity: 53.9% Conservative: 54
Best Local Similarity: 35.5% Mismatches: 117
Query Match: 11.2% Indels: 18
DB: 2 Gaps: 3
US-10-768-158-1 (1-2419) x Q9R1S5_MOUSE (1-298)
QY 36 GCCGAGACCCCGACACCCCGGGGAGTTCGAGAGCAAGTACTTCGAGTTCATCGCGGTG 95
DB 2 AlaGlnAsnProSerAsnMetGluProLeuArgLysProLeuValProValLysGlyLe 21
QY 96 CGGTGCGCGCTCTCCCGCGGAAGATGGAGAGATCGCAACTTCCCGTGGCGGCC 155
DB 22 ProLeuLysTyrPheAlaGluThrMetGluGlnLeuGlnAsnPheThrAlaTrpPro 41
QY 156 AGCAGCTGTGGATCGTCACCTACCCCAAGTCCCGCACCAAGCTTGTGCGAGGTGTC 215
DB 42 AspAspValLeuIleSerThrTyrProLysSerGlyThrAsnTrpMetSerGluMet 61
QY 216 TACTTGTGAGCCAGGCGCTGACCCCGATGAGATCGGCTTGATGAACATCGAGCAG 275
DB 62 AspMetIleTyrGlnGlyGlyLysLeuAspLysCysGlyArgAlaProValTyrAla 81
QY 276 CTCCCGTCTCGGAGTACCCA-----CAGCGCGCTGGACATCATCAAGAA 323
DB 82 IleProPheLeuGluPheSerCysProGlyValProProGlyLeuGluThrLeuLysGlu 101
QY 324 CTGACCTCTCCCGCTCATCAAGAGCCACCTGCGCTTCTACCGCTTCTGCTGACCTC 383
DB 102 ThrProAlaProArgIleIleLysThrHisLeuProLeuSerLeuLeuProGlnSerLeu 121
QY 384 CACAATGAGACTCAAGGTCTATATGCTGTATGCTGCAACCCCAAGGATCTGCTGTGTC 443
DB 122 LeuAspGlnLysIleLysValIleTyrValAlaArgAsnAlaLysAspValValSer 141
QY 444 TATTATCATGTTCCACCGCTCTCTCGGACCATGAGTACCGAGGACCTTTCAAGAAATTC 503
DB 142 TyrTyrAsnPheTyrLysMetAlaLysLeuHisProAspProGlyThrTrpGluSerPhe 161
QY 504 TGGCGGAGTTTATGAATGATAAGCTGGCTACGGCTCTCGCTTGTGACACGTGAGGAG 563
DB 162 LeuGluAsnPheMetAspGlyLysValSerTyrGlySerTyrTyrGlnHisValLysGlu 181
QY 564 TTCTGGGAGCACCAGATGAGTCCGAACTGCTTTTCTCAAGATGAGACATGATCGG 623
DB 182 TrpTrpGluLeuArgArgThrHisProValLeuTyrLeuPheTyrGluAspMetLysGlu 201
QY 624 GACCTGTGTGAGATGGTGGAGCAGCTGGCCAGATTCCTGGGGGTGTCTGTGACAAAGGCC 683
```

```
Db 202 AnProLysArgGluIleLysLysIleLeuGluPheLeuGlyArgSerLeuProGluGlu 221
Qy 684 CAGCTGGAAGCCCTGACGAGCAGTCG-----CACCAGCTGGTGGAGCAGTGTGC 734
Db 222 ThrValAspLeuIleValHisHisThrPheLysLysMetLysGluAsnProMetAla 241
Qy 735 AACGCTGAGGCCCTGCCC-----GTGGGCCGG 761
Db 242 AsnTyThrThrIleProThrGluValMetAspHisThrIleTyThrProPheMetArgLys 261
Qy 762 GGAAGAGTGGCTGGAAGGACATCTTCCCGCTCCATGAATGAGAAGTTTTCACATTG 821
Db 262 GlyThrIleGlyAspTrpLysAsnThrPheThrValAlaGlnSerGluHisPheAspAla 281
Qy 822 GTGTATAACAGAGATGGGAAGTGTGACCTCACGTTT 860
Db 282 HisTyAlaLysLeuMetThrGlyCysAspPheThrPhe 294

RESULT 13
ID ST1A1_MOUSE STANDARD; PRT; 291 AA.
AC P52840;
DT 01-OCT-1996, integrated into UniProtKB/Swiss-Prot.
DT 01-OCT-1996, sequence version 1.
DT 07-FEB-2006, entry version 43.
DE Sulfotransferase 1A1 (EC 2.8.2.1) (Aryl sulfotransferase) (Phenol
DE sulfotransferase) (Sulfokinase) (Phenol/aryl sulfotransferase) (mStp1)
DE (ST1A4).
GN Name=Sult1a1; Synonyms=St1a1, Stp, Stp1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE [mRNA].
RC STRAIN=C57BL/6 X CBA; TISSUE=Liver;
RX MEDLINE=93144346; PubMed=8424956; DOI=10.1016/0167-4781(93)90073-M;
RA Kong A.-N.T., Ma M., Tao D., Yang L.;
RT "Molecular cloning of cDNA encoding the phenol/aryl form of
RT sulfotransferase (mStp1) from mouse liver.";
RL Biochim. Biophys. Acta 1171:315-318(1993).
CC -!- FUNCTION: Catalyzes O-sulfation of phenols, N-O-sulfation of
CC minoxidil and tyrosine esters (By similarity).
CC -!- CATALYTIC ACTIVITY: 3'-phosphoadenylyl sulfate + a phenol =
CC adenosine 3',5'-bisphosphate + an aryl sulfate.
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasm (By similarity).
CC -!- SIMILARITY: Belongs to the sulfotransferase family.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC ENBL; L02331; -, NOT_ANNOTATED_CDS; mRNA.
DR HSSP; P50224; 1CUM.
DR SMR; P52840; 5-291.
DR Ensembl; ENSMUSG00000030711; Mus musculus.
DR MGI; MGI:102896; Sult1a1.
DR GO; GO:0000166; F: nucleotide binding; IDA.
DR GO; GO:0008146; F: sulfotransferase activity; IDA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase 1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
DR Lipid metabolism; Steroid metabolism; Transferase.
FT CHAIN 1 291 Sulfotransferase 1A1.
FT /FTID=PRO_0000085130.
FT PAPS (By similarity).
FT NP_BIND 44 49 PAPS (By similarity).
FT NP_BIND 126 134 PAPS (By similarity).
FT NP_BIND 189 225 PAPS (By similarity).
FT NP_BIND 253 255 PAPS (By similarity).
FT ACT_SITE 104 104 Proton acceptor (By similarity).
SQ SEQUENCE 291 AA; 33974 MW; ADC07187DFAS1D96 CRC64;
```

```
Alignment Scores:
Pred. No.: 8.84e-26 Length: 291
Score: 489.00 Matches: 101
Percent Similarity: 56.0% Conservative: 53
Best Local Similarity: 36.7% Mismatches: 103
Query Match: 11.1% Indels: 18
DB: 1 Gaps: 3

US-10-768-158-1 (1-2419) x ST1A1_MOUSE (1-291)
Qy 90 GCGCTGGCGTCCGCCCTCTGCGCGGGAAGATGGAGAGATCGCCAACTTCCCGGTG 149
Db 13 GlyIleProLeuIleLysTyThrPheAlaGluThrMetGluGlnLeuGluAsnPheThrAla 32
Qy 150 CGGCCACGACGCTGTGGATCGTCACTTACCCCAAGTCCGGCACCAGCTGTGCGAGG 209
Db 33 TrpProAspAspValLeuIleSerThrTyThrProLysSerGlyThrAsnTrpMetSerGlu 52
Qy 210 GTGTCTACTTGTGACCGCGGCTGACCCCGATGAGATCGGCTTGTATGAACATCGAC 269
Db 53 IleMetAspMetIleTyThrGlnGlyGlyLysLeuAspLysCysGlyArgAlaProValTy 72
Qy 270 GAGCAGCTCCCGGCTCTGGAGTACCA-----CAGCGGGCTCGACATCATC 317
Db 73 AlaArgIleProPheLeuGluPheSerCysProGlyValProProGlyLeuGluThrLeu 92
Qy 318 AAGAACTGACCTCTCCCGCTCATCAAGAGCACCTGCGCTACCGCTTTCTGCCCTCT 377
Db 93 LysGluThrProAlaProArgIleIleLysThrHisLeuProLeuSerLeuLeuProGln 112
Qy 378 GACCTCACAAATGAGACTCCAAAGTCACTATATGCTCCCAACCCCAAGATCTGTG 437
Db 113 SerLeuLeuAspGlnLysIleLysValIleTyThrValAlaArgAsnAlaLysAspVal 132
Qy 438 GTGTCTTATTATCAGTTCCACCGCTCTCTGCGGACCATGACTACCGAGGACCTTCAA 497
Db 133 ValSerTyThrAsnPheTyThrLysMetAlaLysLeuHisProAspProGlyThrTrpGlu 152
Qy 498 GAATTTCTCCCGAGGTTTATGAATGATAAGCTGGGCTACGGCTCTCTGTTTGGACACGTG 557
Db 153 SerPheLeuGluAsnPheMetAspGlyLysValSerTyThrGlySerTyThrGlnHisVal 172
Qy 558 CAGAGTTCTGGGAGCACCAGCATGCTGCAACCTGCTTTTCTCAAGTATGAAGACATG 617
Db 173 LysGluTrpTrpGluLeuArgArgThrHisProValLeuTyThrLeuPheTyThrGluAspMet 192
Qy 618 CATCGGACCTGGTGCAGTACGATGGTGGAGCGTGGCCAGATTCTCTGGGGGTCTCTGTGAC 677
Db 193 LysGluAsnProLysArgGluIleLysLysIleLeuGluPheLeuGlyArgSerLeuPro 212
Qy 678 AAGCCCCAGCTGGAAGCCCTGACGAGCAGCTGC-----CACCAGCTGGTGGACAG 728
Db 213 GluGluThrValAspLeuLeuIleValHisHisThrSerPheLysLysMetLysGluAsnPro 232
Qy 729 TGCTGCAACGCTGAGGCCCTGCGCC-----GTG 755
Db 233 MetAlaAsnTyThrThrIleProThrGluValMetAspHisThrIleTyThrProPheMet 252
Qy 756 GCGCGGGAAGAGTTGGCTGTGGGAAGGACATCTTCCACCGTCTCCATCAATGAGAAGTTT 815
Db 253 ArgLysGlyThrIleGlyAspTrpLysAsnThrPheThrValAlaGlnSerGluHisPhe 272
Qy 816 GACTTGTGTATAAACAAGAGATGGGAAGTGTGACCTCACGTTT 860
Db 273 AspAlaHisTyAlaLysLeuMetThrGlyCysAspPheThrPhe 287

RESULT 14
ST1C1_MOUSE
ID ST1C1_MOUSE STANDARD; PRT; 296 AA.
AC Q9D935; O8R210;
DT 13-APR-2004, integrated into UniProtKB/Swiss-Prot.
DT 01-JUN-2001, sequence version 1.
```


DE sulfotransferase 1A3) (Sulfotransferase, monoamine-preferring) (M-BST)
DE (Thermolabile phenol sulfotransferase) (TL-EST) (Placental estrogen
DE sulfotransferase) (Catecholamine-sulfating phenol sulfotransferase)
DE (HAST3).
GN Name=SULT1A3; Synonyms=STM;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE (ISOFORM 1).
RC TISSUE=Brain;
RX MEDLINE=93371391; PubMed=8363392;
RA Zhu X., Veronese M.E., Bernard C.C., Sansom L.N., McManus M.E.;
RT "Identification of two human brain aryl sulfotransferase cDNAs.";
RL Biochem. Biophys. Res. Commun. 195:120-127(1993).
RN [2]
RP NUCLEOTIDE SEQUENCE (ISOFORM 1).
RC TISSUE=Placenta;
RX MEDLINE=94244843; PubMed=8187949; DOI=10.1016/0303-7207(94)90159-7;
RA Bernier F., Lopez-Solache I., Labrie F., Luu-The V.;
RT "Cloning and expression of cDNA encoding human placental estrogen
RT sulfotransferase.";
RL Mol. Cell. Endocrinol. 99:R11-R15(1994).
RN [3]
RP NUCLEOTIDE SEQUENCE (ISOFORM 1).
RX MEDLINE=95100963; PubMed=7802665;
RA Dooley T.P., Probst P., Munroe P.B., Mole S.E., Liu Z., Doggett N.A.;
RT "Genomic organization and DNA sequence of the human catecholamine-
RT sulfating phenol sulfotransferase gene (STM).";
RL Biochem. Biophys. Res. Commun. 205:1325-1332(1994).
RN [4]
RP NUCLEOTIDE SEQUENCE, AND PROTEIN SEQUENCE OF 84-101.
RC TISSUE=Liver;
RX MEDLINE=94161723; PubMed=8117269;
RA Wood T.C., Aksoy I.A., Aksoy S., Weinshilboum R.M.;
RT "Human liver thermolabile phenol sulfotransferase: cDNA cloning,
RT expression and characterization.";
RL Biochem. Biophys. Res. Commun. 198:1119-1127(1994).
RN [5]
RP NUCLEOTIDE SEQUENCE (ISOFORM 1).
RC TISSUE=Platelet;
RX MEDLINE=95209696; PubMed=7695637;
RA Aksoy I.A., Weinshilboum R.M.;
RT "Human thermolabile phenol sulfotransferase gene (STM): molecular
RT cloning and structural characterization.";
RL Biochem. Biophys. Res. Commun. 208:786-795(1995).
RN [6]
RP NUCLEOTIDE SEQUENCE (ISOFORM 1).
RC TISSUE=Blood;
RX MEDLINE=95209704; PubMed=7695643;
RA Jones A.L., Hagen M., Coughtrie M.W.H., Roberts R.C., Giatt H.;
RT "Human platelet phenolsulfotransferases: cDNA cloning, stable
RT expression in V79 cells and identification of a novel allelic variant
RT of the phenol-sulfating form.";
RL Biochem. Biophys. Res. Commun. 208:855-862(1995).
RN [7]
RP NUCLEOTIDE SEQUENCE (ISOFORM 1).
RC TISSUE=Leukocyte;
RX MEDLINE=95050600; PubMed=7961757;
RA Bernier F., Leblanc G., Labrie F., Luu-The V.;
RT "Structure of human estrogen and aryl sulfotransferase gene. Two mRNA
RT species issued from a single gene.";
RL J. Biol. Chem. 269:28200-28205(1994).
RN [8]
RP NUCLEOTIDE SEQUENCE (ISOFORM 1).
RC TISSUE=Liver;
RA Gaeidig A., Grant D.M.;
RT Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
RN [9]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 2).
RX PubMed=14702039; DOI=10.1038/ng1285;

RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Ohayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahari K.,
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shitatori A.,
RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
RA Abe K., Kamiyama K., Matsutani N., Sato K., Tanikawa M., Yamazaki M.,
RA Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
RA Tanai H., Kimata M., Watanabe M., Hiraoaka S., Chiba Y., Ishida S.,
RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J.,
RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O., Nomura Y.,
RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
RA Yushino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA Moriya S., Moriyma H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y.,
RA Ohtani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs.";
RL Nat. Genet. 36:40-45(2004).
RN [10]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).
RC TISSUE=Lung, and Pancreas;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Berge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uesdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.A., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [11]
RP NUCLEOTIDE SEQUENCE OF 139-198.
RC TISSUE=Lymphocyte;
RX MEDLINE=95130098; PubMed=7829089;
RA Aksoy I.A., Callen D.F., Apostolou S., Her C., Weinshilboum R.M.;
RT "Thermolabile phenol sulfotransferase gene (STM): localization to
RT human chromosome 16p11.2.";
RL Genomics 23:275-277(1994).
RN [12]
RP CHARACTERIZATION.
RX MEDLINE=94379981; PubMed=8093002;
RA Veronese M.E., Burgess W., Zhu X., McManus M.E.;
RT "Functional characterization of two human sulfotransferase cDNAs that
RT encode monoamine- and phenol-sulfating forms of phenol
RT sulfotransferase: substrate kinetics, thermal-stability and
RT inhibitor-sensitivity studies.";
RL Biochem. J. 302:497-502(1994).
RN [13]


```
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AJ238391; CAB41460.1; -; mRNA.
DR PIR; JC7282; JC7282.
DR HSP; F50224; ICDM.
DR Ensembl; ENSRNOG00000031833; Rattus norvegicus.
DR RGD; 621064; Sult1c2.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase.
DR ProDom; P0001218; Sulfotransferase; 1.
KW Lysozyme; Transferase.
FT CHAIN 1 296 Sulfotransferase K1.
FT /FTID-PRO_0000085135.
FT NP_BIND 49 54 PAPS (By similarity).
FT NP_BIND 131 139 PAPS (By similarity).
FT NP_BIND 194 230 PAPS (By similarity).
FT NP_BIND 258 260 PAPS (By similarity).
FT ACT_SITE 109 109 Proton acceptor (By similarity).
SQ SEQUENCE 296 AA; 34699 MW; 4A82FAD740EBABF5 CRC64;

Alignment Scores:
Pred. No.: 3 248-24 Length: 296
Score: 467.00 Matches: 100
Percent Similarity: 55.4% Conservative: 58
Best Local Similarity: 35.1% Mismatches: 105
Query Match: 10.6% Indels: 22
DB: 5 Gaps: 5

US-10-768-158-1 (1-2419) x STK1_RAT (1-296)
QY 66 GAGAGCAAGTACTTCGAGTTCATCGCGTTCGCGGCTGCGCGCCCTTCGCGCGGGAAGATG 125
D 10 GlnThrLysLeuLysGluValAlaGlyLeuProLeuGlnAlaProThrValAlaSerTrp 29
QY 126 GAGGAGATCGCAACTTCGCGTGGCGGCGGAGCGAGTGTGATCGTACCTACCCCAAG 185
D 30 SerGlnIleGlnThrPheLysAlaLysProAspLeuLysCysThrTyrProLys 49
QY 186 TCCGSCACGAGTTCGTCAGAGTGTCTACTTGTGGAGCGGCGGCGGCGGCGGCGGAT 245
D 50 SerGlyThrThrTrpIleGlnGluLeuValAspMetIleGluGlnAsnGlyAspValGlu 69
QY 246 GAGATCGGCTTGATGAACATGACGAGCGAGTCCGCGTCTCGGAGTAC-----CCA 296
D 70 LysCysGlnArgThrIleIleGlnHisArgHisProPheIleGluTrpAlaArgProPro 89
QY 297 GAGCGC---GGCTGGACATCATCAAGAACTGACCTCTCCCGGCTCATCAAGAGCCAC 353
D 90 GlnProSerGlyValAspLysAlaAsnAlaMetProAlaProArgIleLeuArgThrHis 109
QY 354 CTGCGCTACCGTTCTCGCCCTCTGACCTCCACATGGAGACTCCCAAGTCTATATG 413
D 110 LeuProThrGlnLeuLeuProSerPheThrThrAsnAsnGlyPheLeuThrVal 129
QY 414 GCTCGCAACCCCAAGGATCTGGTGTCTTATTATCATGTTCCACCGCTCTCTCGGACC 473
D 130 AlaArgAsnAlaLysAspCysMetValSerTyrThrHisPheTyrArgMetSerGlnVal 149
QY 474 ATGAGCTACCGAGGACCTTTCAAGATTTCTCGCGAGGTTTAAATGATAGTCTGGC 533
D 150 LeuProAspProGlyThrTrpAsnGluThrPheGluThrPheIleAsnGlyLysValSer 169
QY 534 TACGCTCTGTTGACAGCTGAGGAGTCTCGGAGGACCGGACCGGAGTCTCGAAGCTG 593
D 170 TrpGlySerTrpPheAspHisValLysGlyTrpTrpGluIleArgAspArgTyrGlnIle 189
QY 594 CTTTTCCTCAAGTGAAGACATGATCGGAGCTGGTGGACGATGGGAGGAGTCTGGCC 653
D 190 LeuPheLeuPheTyrGluAspValLysArgAspProLysArgGluIleGlnLysValMet 209
QY 654 AGATTCTCGGGGTGTCC-----TGTGCAAGGCCAGCTGGAAGCCCTG 698
```

```
Db 210 GlnPheMetGlyLysAsnLeuAspGluValValAspLysIleValLeuGluThrSer 229
QY 699 ACGAGCAGCTGCCACCACTGGTGCACAGCTGCACAGCTGCAGCGCTGCCCTCC-- 752
D 230 PheGluLysMetLysGlu-----AsnProMetThrAsnArgSerThrValProLysSer 247
QY 753 -----GTGGGCGGGGAGAGAGTGGGCTGTGAAGAC 785
D 248 ValLeuAspGlnSerIleSerProPheMetArgLysGlyThrValGlyAspTrpLysAsn 267
QY 786 ATCTTCACCGTCTCCATCAATGAGAGTTGACTTGGTATTAACAGAACATGGGAAG 845
D 268 HisPheThrValAlaGlnAsnAspArgPheAspGluIleTyrLysGlnLysMetGly 287
QY 846 TGTGACCTTCAGCTTT 860
D 288 ThrSerLeuAsnPhe 292

RESULT 18
Q3ZAV3 RAT PRELIMINARY; PRT; 296 AA.
AC Q3ZAV3;
DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.
DT 22-NOV-2005, sequence version 1.
DE 07-FEB-2006, entry version 3.
DE Sulfotransferase family, cytosolic, 1C, member 2.
GN Name=Sult1c2;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Berge J.G.,
RA Klausner R.D., Collins L.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RG NIH MGC Project;
RL Submitted (AUG-2005) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BC103636; AA103637.1; -; mRNA.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
KW Transferase.
SQ SEQUENCE 296 AA; 34699 MW; 4A82FAD740EBABF5 CRC64;

Alignment Scores:
```

```
Pred. No.: 3.24e-24 Length: 296
Score: 467.00 Matches: 100
Percent Similarity: 55.4% Conservat: 58
Best Local Similarity: 35.4% Mismatches: 105
Query Match: 22 Indels: 22
DB: 2 Gaps: 5

US-10-768-158-1 (1-2419) x Q3ZAV3_RAT (1-296)
Qy 66 GAGAGCAAGTACTCGAGTTCATGGCTGGCGGTGCGGCTCTGCGCGGGAAGATG 125
Db 10 GlnThrLysLeuLysGluValAlaGlyIleProLeuGlnAlaProThrValAsnTrp 29
Qy 126 GAGGAGATCGCAACTTCCCGGTGGCGCCAGCGAGTGGATCGTCACCTACCCCAAG 185
Db 30 SerGlnIleGlnThrPheLysAlaLysProAspAspLeuLeuIleCysThrTrpProLys 49
Qy 186 TCCGGCACCAGCTTGCTCAGAGAGTGTCTACTTGTGTAGCCAGCGGTGACCCCGAT 245
Db 50 SerGlyThrTrpIleGlnGluIleValAspMetIleGluGlnAsnGlyAspValGlu 69
Qy 246 GAGATCGCTTGATGAACATGACGAGAGCTCCCGGTCTGGAGTAC-----CCA 296
Db 70 LysCysGlnArgThrIleIleGlnHisArgHisProPheIleGluTrpAlaArgPro 89
Qy 297 CAGCG---GGCTGGACATCATCAAGAACTGACCTCTCCCGCTCATCAAGAGCCAC 353
Db 90 GlnProSerGlyValAspLysAlaAsnAlaMetProAlaProArgIleLeuArgThrHis 109
Qy 354 CTGCGCTACCGCTTCTCCCTCTGACCTCCACAAATGGAGACTCCAAAGTCTATATG 413
Db 110 LeuProThrGlnLeuLeuProProSerPheTrpThrAsnAsnCysLysPheLeuTyVal 129
Qy 414 GCTGCAACCCCAAGATCTGGTGTCTTATATCATGTTCCACCGCTCTCTCGGAGC 473
Db 130 AlaArgAsnAlaLysAspCysMetValSerTyTrpHisPheTyArgMetSerGlnVal 149
Qy 474 ATGAGTACCAGGACCTTTCAGAAATTCGCGAGGTTTATGAATGATAAGTGGC 533
Db 150 LeuProAspProGlyThrTrpAsnGlnTyPheGluThrPheIleAsnGlyLysValSer 169
Qy 534 TACGCTCTCTGTTGACACGTGACGAGTCTTGGGAGCACCGCATGGACTCGAACGTG 593
Db 170 TrpGlySerTrpPheAspHisValLysGlyTrpTrpGluIleArgAspArgTyGlnIle 189
Qy 594 CTTTTTCAAGTATGAAGATCATGCTCGGACCTGTGTGACGATGGTGGAGCAGCTGGCC 653
Db 190 LeuPheLeuPheTyArgGluAspValLysArgAspProLysArgGluIleGlnLysValMet 209
Qy 654 AGATTCTCTGGGGTGTCC-----TGTGACAAGGCCACGCTGGAAGCCCTG 698
Db 210 GlnPheMetGlyLysAsnLeuAspGluValValAspLysIleValLeuGluThrSer 229
Qy 699 ACGGAGCACTGCCACCACTGGTGACGACCTGTGCAACGCTGAGCGCCCTGCC----- 752
Db 230 PheGluLysMetLysGlu-----AsnProMetThrAsnArgSerThrValProLysSer 247
Qy 753 -----GTGGCGCGGGAAGAGTGTGGCTGTGGAAGAC 785
Db 248 ValLeuAspGlnSerIleSerProPheMetArgLysGlyThrValGlyAspTrpLysAsn 267
Qy 786 ATCTTTCACCGTCTCCATCAATGAGAAGTTTGCATTTGGTGTATAAACAGAAGATGGAAAG 845
Db 268 HisPheThrValAlaGlnAsnAspArgPheAspGluIleTyTrpLysGlnLysMetGlyGly 87
Qy 846 TGTGACCTCAGCTTT 860
Db 288 ThrSerLeuAsnPhe 292

RESULT 19
Q642G8_RAT PRELIMINARY; PRT; 296 AA.
ID Q642G8_RAT
AC Q642G8;
```

25-OCT-2004, integrated into UniprotKB/TrEMBL.
25-OCT-2004, sequence version 1.
07-FEB-2006, entry version 11.
Sulfotransferase family, cytosolic, 1C, member 1 (Predicted).
GN Name=Sultici predicted;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
[1]
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Kidney;
RC MEDLINE=2238257; PubMed=12477932; DOI=10.1073/pnas.242603899;
EX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uesdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield Y.S.N., Krzywinski M.I., Skaleka U., Smalls D.E.,
RA Schemm A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences".
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[2]
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Kidney;
RC NIH MGC Project;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the Uniprot Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License

CC EMBL: BC081691; AAH81691.1; -; mRNA.
DR SMR: Q642G8; 12-296.
DR Ensembl: ENSRNOG000000031833; Rattus norvegicus.
DR GO: 0008146; F:sulfotransferase activity; IEA.
DR GO: 0016740; F:transferase activity; IEA.
DR InterPro: IPR000863; Sulfotransferase.
DR Pfam: PF00685; Sulfotransferase 1; 1.
DR ProDom: PD001218; Sulfotransferase; 1.
KW Transferase.
SQ SEQUENCE 296 AA; 34859 MW; FB14A27E39BE847F CRC64;
Alignment Scores:
Pred. No.: 3.81e-24 Length: 296
Score: 466.00 Matches: 101
Percent Similarity: 55.1% Conservat: 56
Best Local Similarity: 35.4% Mismatches: 106
Query Match: 22 Indels: 22
DB: 2 Gaps: 5
US-10-768-158-1 (1-2419) x Q642G8_RAT (1-296)
Qy 66 GAGAGCAAGTACTCGAGTTCATGGCTGGCGGTGCGGCTCTGCGCGGGAAGATG 125
Db 10 GlnThrLysLeuLysGluValAlaGlyIleProLeuGlnAlaProThrValAsnTrp 29
Qy 126 GAGGAGATCGCAACTTCCCGGTGGCGCCAGCGAGTGGATCGTCACCTACCCCAAG 185
Db 30 SerGlnIleGlnThrPheLysAlaLysProAspAspLeuLeuIleCysThrTrpProLys 49
Qy 186 TCCGGCACCAGCTTGCTCAGAGAGTGTCTACTTGTGTAGCCAGCGGTGACCCCGAT 245
Db 50 SerGlyThrTrpIleGlnGluIleValAspMetIleGluGlnAsnGlyAspValGlu 69
Qy 246 GAGATCGCTTGATGAACATGACGAGAGCTCCCGGTCTGGAGTAC-----CCA 296
Db 70 LysCysGlnArgThrIleIleGlnHisArgHisProPheIleGluTrpAlaArgPro 89
Qy 297 CAGCG---GGCTGGACATCATCAAGAACTGACCTCTCCCGCTCATCAAGAGCCAC 353
Db 90 GlnProSerGlyValAspLysAlaAsnAlaMetProAlaProArgIleLeuArgThrHis 109
Qy 354 CTGCGCTACCGCTTCTCCCTCTGACCTCCACAAATGGAGACTCCAAAGTCTATATG 413
Db 110 LeuProThrGlnLeuLeuProProSerPheTrpThrAsnAsnCysLysPheLeuTyVal 129
Qy 414 GCTGCAACCCCAAGATCTGGTGTCTTATATCATGTTCCACCGCTCTCTCGGAGC 473
Db 130 AlaArgAsnAlaLysAspCysMetValSerTyTrpHisPheTyArgMetSerGlnVal 149
Qy 474 ATGAGTACCAGGACCTTTCAGAAATTCGCGAGGTTTATGAATGATAAGTGGC 533
Db 150 LeuProAspProGlyThrTrpAsnGlnTyPheGluThrPheIleAsnGlyLysValSer 169
Qy 534 TACGCTCTCTGTTGACACGTGACGAGTCTTGGGAGCACCGCATGGACTCGAACGTG 593
Db 170 TrpGlySerTrpPheAspHisValLysGlyTrpTrpGluIleArgAspArgTyGlnIle 189
Qy 594 CTTTTTCAAGTATGAAGATCATGCTCGGACCTGTGTGACGATGGTGGAGCAGCTGGCC 653
Db 190 LeuPheLeuPheTyArgGluAspValLysArgAspProLysArgGluIleGlnLysValMet 209
Qy 654 AGATTCTCTGGGGTGTCC-----TGTGACAAGGCCACGCTGGAAGCCCTG 698
Db 210 GlnPheMetGlyLysAsnLeuAspGluValValAspLysIleValLeuGluThrSer 229
Qy 699 ACGGAGCACTGCCACCACTGGTGACGACCTGTGCAACGCTGAGCGCCCTGCC----- 752
Db 230 PheGluLysMetLysGlu-----AsnProMetThrAsnArgSerThrValProLysSer 247
Qy 753 -----GTGGCGCGGGAAGAGTGTGGCTGTGGAAGAC 785
Db 248 ValLeuAspGlnSerIleSerProPheMetArgLysGlyThrValGlyAspTrpLysAsn 267
Qy 786 ATCTTTCACCGTCTCCATCAATGAGAAGTTTGCATTTGGTGTATAAACAGAAGATGGAAAG 845
Db 268 HisPheThrValAlaGlnAsnAspArgPheAspGluIleTyTrpLysGlnLysMetGlyGly 87
Qy 846 TGTGACCTCAGCTTT 860
Db 288 ThrSerLeuAsnPhe 292

Db	150	LeuProAsnProGlyThrTrpAsnGluTyrPheGluThrPheIleAsnGlyLysValSer	169
Qy	534	TACGGCTCTCGTTTGAGCACGTCGAGGAGTCTGGGAGCACCGACTGGACTGCAACGTG	593
Db	170	TrpGlySerTrpPheAspHisValLysGlyTrpTrpGluMetArgAspArgTyrGlnIle	189
Qy	594	CTTTTCTCAAGTATGAACACATGCATCGGACCTGGTGCAGTCATGTGTGGACAGCTGGCC	653
Db	190	LeuPheLeuPheTyrGluAspMetLysArgAspProLysArgGluIleGlnLysValMet	209
Qy	654	AGATTCTCGGGGTGTC-----TGTGACAAAGCCACGCTGGAGCCCTG	698
Db	210	GlnPheMetGlyLysAsnLeuAspGluGluValValAspLysIleValLeuGluThrSer	229
Qy	699	ACGGAGCACTGCCACCAG-----CTGGTG	722
Db	230	PheGluLysMetLysAspAsnProLeuThrAsnPheSerThrIleProLysThrIleMet	249
Qy	723	GACCACTGTCGAACGCTGAGCCCTGCCCTGGCGCGGGAAGAGTTGGCTGTGGAAG	782
Db	250	AspGln-----SerIleSerLeuPheMetArgLysGlyIleValGlyAspTrpLys	266
Qy	783	GACATCTTCCCGTCTCCATGAATGAGAAGTTTGACTTGTGTATAAACAGAGATGGGA	842
Db	267	AsnHisPheThrValAlaGlnAsnGluArgPheAspGluIleTyrGluGlnLysLeuAsp	286
Qy	843	AAGTGTGACTCAGCTTT	860
Db	287	GlyThrSerLeuAsnPhe	292
RESULT 21			
Q2TAB3_HUMAN			
ID_Q2TAB3_HUMAN	PRELIMINARY;	PRT;	295 AA.
AC_Q2TAB3;			
DT_24-JAN-2006,	integrated into UniProtKB/TrEMBL.		
DT_24-JAN-2006,	sequence version 1.		
DT_07-FEB-2006,	entry version 2.		
DE_Sulfotransferase family,	cytosolic, 1A, phenol-preferring, member 3.		
GN_Name=SU1T1A3;			
OS_Homo sapiens (Human).			
OC_Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC_Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;			
OC_Homo.			
OX_NCBI_TaxID=9606;			
RP[1]			
NUCLEOTIDE SEQUENCE.			
RC_TSSUE=Brain;			
RX_MEDLINE=22388257;	PubMed=12477932; DOI=10.1073/pnas.242603899;		
RA_Krausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA_Klauniger R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA_Altschul S.F., Zebberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA_Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,			
RA_Diatchenko L., Marxina K., Farmer A.A., Rubin G.M., Hong L.,			
RA_Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA_Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,			
RA_Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,			
RA_Boak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA_Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyar S.W.,			
RA_Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA_Fahy J., Heaton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,			
RA_Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA_Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RA_Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,			
RA_Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,			
RA_Schnurch A., Schein J.E., Jones S.J.M., Marra M.A.;			
RT"Generation and initial analysis of more than 15,000 full-length human			
RT and mouse cDNA sequences."			
PL_Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).			
RP[2]			
NUCLEOTIDE SEQUENCE.			
RC_TISSUE=Brain;			
RG_NIH MGC Project;			

RA Heaton M.P., Laegreid W.W., Rohrer G.A., Chitko-McKown C.G.,
 RA Pettea G., Holt I., Karamycheva S., Liang F., Quackenbush J.,
 RA Keele J.W.;
 RT "Sequence evaluation of four pooled-tissue normalized bovine cDNA
 RT libraries and construction of a gene index for cattle.";
 RL Genome Res. 11:626-630(2001).
 RN (2)
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Poolled;
 RA Hathay G.P., Sonstegard T.S., Van Tassel C.P., Clawson M.L.,
 RA Heaton M.P., Keele J.W., Snelling W.M., Weidmann R.T., Smith T.P.L.;
 RT "Sequencing and analysis of Bos taurus full-length insert cDNA
 RT clones";
 RL Submitted (MAR-2005) to the ENBL/GenBank/DBJ databases.
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC (2)
 DR EMBL; BT021839; AAX46686.1; -; mRNA.
 DR SMR; Q58CV8; 11-262.
 DR Ensembl; ENSBTAG00000016032; Bos taurus.
 DR GO; GO:0008146; F:sulfotransferase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR InterPro; IPR000863; Sulfotransferase.
 DR Pfam; PF00685; Sulfotransfer_1; 1.
 KW Transferase.
 KW SEQUENCE 298 AA; 34914 MW; E597E9EA1032E37E CRC64;
 Alignment Scores:
 Pred. No.: 6,23e-24 Length: 298
 Score: 463.00 Matches: 98
 Percent Similarity: 56.1% Conservative: 58
 Best Local Similarity: 35.3% Mismatches: 102
 Query Match: 10.5% Indels: 20
 DB: 2 Gaps: 5

US-10-768-158-1 (1-2419) x Q58CV8_BOVIN (1-298)

QY 90 GCGTGGCGTCCGCG---CCCTTCTGCGCGGGAAGATGGAGGAGATCGCACTTCCG 146
 DQ 19 GlytLeuGlnProLysProThrCys---AspThrTrpAspGlnileTrpSerPheGln 37
 QY 147 GTGGCGCCGACGAGTGTGTGATCGTACCTACCCCAAGTCCGCGACAGCTTGTGCGAG 206
 DQ 38 AlaArgProAspAspLeuLeuLeuSerThrTyrProLysAlaGlyThrThrTrpThrGln 57
 QY 207 GAGTGTGTCTACTTGGTGAGCGAGCGCTGACCCCGATGAGATCGGCTTGATGAACATC 266
 DQ 58 GluLeuValAspLeuLeuGlnAsnGlyAspValAsnGlnSerGlnArgAlaProThr 77
 QY 267 GACGAGCAGCTCCCGGTCTCGAGTACCCACAGCGG-----GGCTGGACATC 314
 DQ 78 HisGluArgPheProPheIleGluTrpIleProSerLeuGlySerGlyLeuGluGln 97
 QY 315 ATCAAGGAACGACTCTCCCGCTCATCAAGACACCTGCGCTTCTTCTGCGC 374
 DQ 98 AlaAsnAlaMetAlaSerProArgMetLeuLysThrHisLeuProPheHisLeuLeuPro 117
 QY 375 TCTGACCTCCCAATGGAGACTCCAAAGTCTATATATGGTTCGCAACCCCAAGATCTG 434
 DQ 118 ProSerPheLeuGluAsnCysLysMetIleTyrValAlaArgAsnProLysAspAsn 137
 QY 435 GTGGTGTCTTATTATCAGTCTCCCGCTCTCTCGGACCATGAGTACCGGAGGACCTTT 494
 DQ 138 MetValSerTyrTyrHisPheHisArgMetAsnArgAsnLeuProAlaProGlyThrTrp 157
 QY 495 CAAGAATCTCGCGAGGTCTTATGATCATAGCTGGCTCGGCTCTGTTTGTAGCAC 554
 DQ 158 GluGluTyrPheGluSerPheLeuAlaGlyLysValCysTrpGlySerTrpTyrAspHis 177
 QY 555 GTGCAGGAGTCTTCGGGACACCGCATGAGTACGACGTCGCTTTTCTCAAGTATGAAGAC 614
 DQ 178 ValLysGlyTyrTrpTrpGlnAlaLysAspGlnHisArgIleLeuTyrLeuPheTyrGluAsp 197

QY 615 ATGCATCGGACCTGGTGACGATGGTGGAGCAGCTGGCCAGATTCCTCGGGGTCTCTGT 674
 DQ 198 MetLysGluAsnProLysHisGluIleGlnLysLeuAlaGluPheIleGlyLysLeu 217
 QY 675 GACAAAGCCCGACGCTGGAAGCCCTGACGAGCAGTCCGACGAGCTGGTGGT-----GAC 725
 DQ 218 AspAspLysValLeuAspLysIleValAspHisThrSerPheSerValMetLysGlnAsn 237
 QY 726 CAGTGTGCAACGCTGAGGCGCTGGCC-----752
 DQ 238 ProMetAlaAsnTyrThrSerIleProAsnGluTyrMetAsnGlnLeuIleSerProPhe 257
 QY 753 GTGGCGCGGAGAGAGTGGCTGTGGAAGACATCTTCACCGTCTCCATGATGAGAAG 812
 DQ 258 MetArgLysGlyValIleGlyAspTrpLysAsnHisPheThrValAlaGlnAsnGluArg 277
 QY 813 TTTGACTGTGTATAAACAGAGATGGAAAGTGTGACCTCAGCTTGTGACTTT 866
 DQ 278 PheAspAspAspTyrArgLysAsnMetAlaAspThrThrLeuThrLeuHisPhe 295

RESULT 24
 ST1A1 RAT
 ID ST1A1 RAT STANDARD; PRT; 291 AA.
 AC P17986;
 DT 01-NOV-1990, integrated into UniProtKB/Swiss-Prot.
 DT 01-NOV-1990, sequence version 55.
 DE Sulfotransferase 1A1 (EC 2.8.2.1) (Aryl sulfotransferase) (Phenol
 DE sulfotransferase) (PST-1) (Sulfokinase) (Aryl sulfotransferase IV)
 DE (ASTIV) (Tyrosine-ester sulfotransferase) (Minoxidil
 DE sulfotransferase).
 GN Name=Stial; Synonyms=Stial;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muroidae; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 [1]
 NP NUCLEOTIDE SEQUENCE [MRNA].
 RP STRAIN=Sprague-Dawley; TISSUE=Liver;
 RC MEDLINE=90326537; PubMed=2374726;
 RX Ozawa S., Nagata K., Gong D., Yamazoe Y., Kato R.;
 RT "Nucleotide sequence of a full-length cDNA (PST-1) for aryl
 RT sulfotransferase from rat liver.";
 RL Nucleic Acids Res. 18:4001-4001(1990).
 [2]
 NP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
 RP STRAIN=Sprague-Dawley; TISSUE=Liver;
 RC MEDLINE=94131305; PubMed=8299966; DOI=10.1016/0378-1119(93)90028-2;
 RX Khan A.S., Taylor B.R., Chung K., Echeredge J., Gonzales R.,
 RA Ringer D.P.;
 RT "Genomic structure of rat liver aryl sulfotransferase IV-encoding
 RT gene.";
 RL Gene 137:321-326(1993).
 [3]
 NP NUCLEOTIDE SEQUENCE OF 7-291.
 RX MEDLINE=92379794; PubMed=1511441;
 RA Yerokun T., Etheredge J.L., Norton T.R., Carter H.A., Chung K.H.,
 RA Birckbichler P.J., Ringer D.P.;
 RT "Characterization of a complementary DNA for rat liver aryl
 RT sulfotransferase IV and use in evaluating the hepatic gene transcript
 RT levels of rats at various stages of 2-acetylaminofluorene-induced
 RT hepatocarcinogenesis.";
 RL Cancer Res. 52:4779-4786(1992).
 [4]
 RN CHARACTERIZATION.
 RP TISSUE=Liver;
 RC MEDLINE=93191709; PubMed=8447833;
 RX Cruickshank D., Sansom L.N., Veronese M.E., Mojarrahi B.,
 RA McManus M.E., Zhu X.;
 RT "cDNA expression studies of rat liver aryl sulphotransferase.";
 RL Biochem. Biophys. Res. Commun. 191:295-301(1993).


```
RT human phenol sulfotransferase STP2 gene." ;
RL Genomics 40:242-246(1997).
RN [7]
RP CHARACTERIZATION.
RX MEDLINE=95196680; PubMed=789867;
RA Yamazoe Y, Ozawa S, Nagata K., Gong D.-W., Kato R.;
RT "Characterization and expression of hepatic sulfotransferase involved
RL in the metabolism of N-substituted aryl compounds." ;
RL Environ. Health Perspect. 102:99-103(1994).
RN [8]
RP VARIANT ASN-235.
RX MEDLINE=20222641; PubMed=10762004;
RX DOI=10.1097/0000571-200003000-00008;
RA Engelke C.E., Meinel W., Boeing H., Glatt H.;
RT "Association between functional genetic polymorphisms of human
RL sulfotransferases 1A1 and 1A2." ;
RL Pharmacogenetics 10:163-169(2000).
CC -!- FUNCTION: Catalyzes the sulfate conjugation of catecholamines,
CC phenolic drugs and neurotransmitters. Is also responsible for the
CC sulfation and activation of minoxidil. Mediates the metabolic
CC activation of carcinogenic N-hydroxyarylamines to DNA binding
CC products and could so participate as modulating factor of cancer
CC risk.
CC
CC -!- CATALYTIC ACTIVITY: 3'-phosphoadenylyl sulfate + a phenol =
CC adenosine 3',5'-bisphosphate + an aryl sulfate.
CC
CC -!- SUBUNIT: Homodimer (By similarity).
CC
CC -!- SUBCELLULAR LOCATION: Cytoplasm.
CC
CC -!- SIMILARITY: Belongs to the sulfotransferase family.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; U28170; AAB09659.1; -; mRNA.
DR EMBL; U28169; AAB09658.1; -; mRNA.
DR EMBL; X78282; CAA55088.1; -; mRNA.
DR EMBL; U34804; AAB09758.1; -; Genomic DNA.
DR EMBL; U72202; AAB08970.1; -; Genomic DNA.
DR EMBL; U72196; AAB08970.1; JOINED; Genomic DNA.
DR EMBL; U72197; AAB08970.1; JOINED; Genomic DNA.
DR EMBL; U72198; AAB08970.1; JOINED; Genomic DNA.
DR EMBL; U72199; AAB08970.1; JOINED; Genomic DNA.
DR EMBL; U72200; AAB08970.1; JOINED; Genomic DNA.
DR EMBL; U72201; AAB08970.1; JOINED; Genomic DNA.
DR EMBL; U72203; AAB08970.1; JOINED; Genomic DNA.
DR EMBL; U76619; AAB18753.1; -; Genomic DNA.
DR EMBL; U33886; AAC51149.1; -; Genomic DNA.
DR PIR; G01843; G01843.
DR PIR; JC5249; JC5249.
DR HSSP; P50224; 1CJM.
DR SMR; P50226; 8-295.
DR
DR Ensembl; ENSG00000197165; Homo sapiens.
DR HGNC; HGNC:11454; SULT1A2.
DR
DR MIM; 601292; gene.
DR
DR GO; GO:0008146; F:sulfotransferase activity; TAS.
DR GO; GO:0009309; P:amine biosynthesis; TAS.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase_1.
DR ProDom; PD001218; Sulfotransferase; 1.
DR Catecholamine metabolism; Lipid metabolism; Polymorphism;
KW Steroid metabolism; Transferase.
KW
CHAIN 1 295 Sulfotransferase 1A2.
/FTID=PRO_0000085128.
FT NP_BIND 48 53 PAPS (By similarity).
FT NP_BIND 130 138 PAPS (By similarity).
FT NP_BIND 193 229 PAPS (By similarity).
FT NP_BIND 257 259 PAPS (By similarity).
FT ACT_SITE 108 108 Proton acceptor (By similarity).
FT VARIANT 7 7 T -> I.
FT
FT VARIANT 235 235 /FTID=VAR_007426.
FT
FT VARIANT 19 19 /FTID=VAR_007427.
FT CONFLICT 19 19 P -> L (in Ref. 2).
FT CONFLICT 290 290 S -> T (in Ref. 6).
FT SEQUENCE 295 AA; 34286 MW; FAD88E348B49BESA CRC64;
SQ
Alignment Scores:
Pred. No.: 1.95e-23 Length: 295
Score: 456.00 Matches: 94
Percent Similarity: 54.5% Conservative: 58
Best Local Similarity: 33.7% Mismatches: 101
Query Match: 10.4% Indels: 26
DB: 1 Gaps: 4
US-10-768-158-1 (1-2419) x ST1A2_HUMAN (1-295)
QY 90 GGCGTGGCGGTCGCCCTTCTGCGCGGGAAGATGGAGGAGATCGCAACTTCCCGTG 149
Db ||||| ||| : : : : : |||||
QY 17 GlyValProLeuLeuLysTyrPheAlaGluLeuGlyProLeuGlnSerPheGlnAla 36
Db ||||| ||||| : : : : : |||||
QY 150 CGGCCAGCGAGTGTGTGATCGTCACTTACCCAGTCCGCGCACGAGTCTTCCGAGG 209
Db ||||| ||||| : : : : : |||||
QY 37 ArgProAspLeuLeuLeuSerThrTyrProLysSerGlyThrTyrPheValSerGln 56
Db ||||| ||||| : : : : : |||||
QY 210 GTGTCTACTTGTGGAGCGGCGTACCCCGATGAGATCGGCTTGATGAACATCGAC 269
Db : : : : : ||||| : : : : :
QY 57 IleLeuAspMetIleTyrGlnGlyGlyAspLeuGlyLysCysHisArgAlaProIlePhe 76
Db : : : : : ||||| : : : : :
QY 270 GAGCAGCTCCCGGTCCTGGAGTACCCACAGCGG-----GGCTGGACATCATC 317
Db : : : : : ||||| : : : : :
QY 77 MetArgValProPheLeuGluPheLysValProGlyIleProSerGlyMetGluThrLeu 96
Db : : : : : ||||| : : : : :
QY 318 AAGGAAGTGCCTCTCCCGCTCATCAAGAGCCACCTGCTCCCTACCGCTTTCTGCCCTCT 377
Db ||||| ||||| : : : : : |||||
QY 97 LysAsnThrProAlaProArgLeuLeuLysThrHisLeuProLeuAlaLeuLeuProGln 116
Db : : : : : ||||| : : : : :
QY 378 GACTCCACATGAGACTCCAGGTCACTATATGCTCGCAACCCCAAGATCTGTGTG 437
Db ||||| : : : : : ||||| : : : : :
QY 117 ThrLeuLeuAspGlnLysValLysValValTyrValAlaArgAsnAlaLysAspValAla 136
Db : : : : : ||||| : : : : :
QY 438 GTGTCTTATTATCAGTTCACCGCTCTCTCGGACCATGAGTACCGAGGCGACCTTTCAA 497
Db ||||| ||||| : : : : : |||||
QY 137 ValSerTyrTyrHisPheTyrHisMetAlaLysValTyrProHisProGlyThrTrpGlu 156
Db : : : : : ||||| : : : : :
QY 498 GAATTCCTCGCGAGGTTTATGAATGAATAAGCTGCGGTACGGCTCTGTTTGAGCACGTG 557
Db ||||| ||||| : : : : : |||||
QY 157 SerPheLeuGluLysPheMetAlaGlyGluValSerTyrGlySerTyrPheGlnHisVal 176
Db : : : : : ||||| : : : : :
QY 558 CAGGAGTCTTGGGAGACCGCATGCGAAGCTGCTGTTTCTCAAGATATGAAGACATG 617
Db ||||| : : : : : ||||| : : : : :
QY 177 GlnGluTrpTrpGluLeuSerArgThrHisProValLeuTyrLeuPheTyrGluAspMet 196
Db : : : : : ||||| : : : : :
QY 618 CATCGGACCTGGTGACGATGGTGAGCAGTCCAGATTCTCTGGGGGTCTCTGTGAC 677
Db : : : : : ||||| : : : : :
QY 197 LysGluAsnProLysArgGluIleGlnLysIleLeuGluPheValGlyArgSerLeuPro 216
Db : : : : : ||||| : : : : :
QY 678 AAGGCCAGCTGGAAGCCCTGACCGAGCATCTGC----- 710
Db : : : : : |||||
QY 217 GluGluThrValAspLeuMetValGluHisThrSerPheLysGluMetLysLysThrPro 236
Db : : : : : |||||
QY 711 -----CACAGCTGGTGGACCGACAGTCTGCAACGCTGAGGCC 746
Db : : : : : ||||| : : : : :
QY 237 MetThrAsnTyrThrThrValArgGluPheMetAspHisSerIleSer----- 253
Db : : : : : ||||| : : : : :
QY 747 CTGCCC---GTGGCGCGGGAAGATGGGTGGGAAGGACATCTTCCCGTCTCCATG 803
Db ||||| : : : : : ||||| : : : : :
QY 254 ---ProPheMetArgLysGlyMetAlaGlyAspTrpLysThrThrPheThrValAlaGln 272
Db : : : : : ||||| : : : : :
QY 804 AATGAGAAGTTTGTACTTGGTGTATAAACAAGAGATGGAAAGTGTGACCTCACCTTT 860
Db ||||| : : : : : ||||| : : : : :
QY 273 AsnGluArgPheAspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe 291
Db : : : : : ||||| : : : : :
RESULT 27
Q6Y0X5.PIG
ID Q6Y0X5.PIG PRELIMINARY; PRT; 295 AA.
AC Q6Y0X5;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
```



```

DT 07-FEB-2006, entry version 9.
DE Phenol sulfotransferase.
OS Sus scrofa (pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
OC Sub.
OX NCBI_TaxID=9823;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RX PubMed=15014971; DOI=10.1007/s00335-002-2318-4;
RA Lin Z., Lou Y., Squires J.E.;
RT "Molecular cloning and functional analysis of porcine SULT1A1 gene and
RT its variant: a single mutation SULT1A1 causes a significant decrease
RT in sulfation activity.";
RL Mamm. Genome 15:218-226(2004).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AY193893; AAP03092.1; -; mRNA.
DR HSRP; P49891; 1AQU.
DR SMR; Q6Y0X5; 9-295.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase_1.
DR ProDom; PD001218; Sulfotransferase; 1.
DR Transferase.
SQ SEQUENCE 295 AA; 34411 MW; 1721E635F6B94EE0 CRC64;

Alignment Scores:
Pred. No.: 1.95e-23 Length: 295
Score: 456.00 Matches: 98
Percent Similarity: 54.2% Conservative: 51
Best Local Similarity: 35.6% Mismatches: 108
Query Match: 10.4% Indels: 18
DB: 2 Gaps: 3

US-10-768-158-1 (1-2419) x Q6Y0X5_PIG (1-295)
QY 90 GCGCTGCGCTGCGCCCTCTGCGCGGGAAGATGGAGGATCGCCAACTCCCGGTG 149
DB 17 GlyValProLeuIleLeuTyrPheAlaGluAlaLeuGlyProLeuGluSerPheGlnAla 36
QY 150 CGGCCAGCAGCGTGTGGATCGTCACCTACCCCAAGTCCGACAGCTTCTCGAGAG 209
DB 37 TrpProAspAspValLeuIleSerThrTyrProLysSerGlyThrThrTrpValSerGlu 56
QY 210 GTGTCTACTTGTGAGCCAGGCGCTGACCCCGATGAGATCGGCTTGATGAACATCGAC 269
DB 57 IleLeuAspLeuIleTyrGlnGlyAspLeuGlnLysCysGlnArgAlaProIlePhe 76
QY 270 GAGCAGCTCCCGCTCGAGTACCCACAGCCG-----GGCTCGACATCATC 317
DB 77 ValArgValProPheLeuGluPheLysIleProArgCysProThrGlyPheGluLeuLeu 96
QY 318 AAGAACTGACCTCTCCCGCTCATCAAGACCCAGCCCTGCGCTTCTGCCCTCT 377
DB 97 LysAspThrProAlaProArgLeuLeuLysThrHisLeuProLeuThrLeuLeuProGln 116
QY 378 GACTCCACATGGAGACTCCAAAGTCTATATGGCTCGCAACCCCAAGATCTGGTG 437
DB 117 ThrLeuLeuAspGlnLysValIleValValTyrValAlaIleArgAsnAlaLysAspValAla 136
QY 438 GTGTCTTATTATCAGTTCCACCGCTCTCTCGCGGACCATGAGCTACCGAGGACCTTTCAA 497
DB 137 ValSerTyrTyrHisPheTyrArgMetAlaLysValTyrProAsnProGlyThrTrpAsp 156
QY 498 GAATTCTCCCGAGGTTTATGAATGATTAAGCTGGGCTACGGCTCTGTTGGTGGACAGCTG 557
DB 157 SerPheLeuGluAspPheMetAlaGlyGluValSerTyrGlySerTyrTrpTyrGlnHisVal 176

```

```

QY 558 CAGGAGTTCTGGGAGCACCAGCATGCAAGCGTCTTTTCTCAAGTATGAAGACATG 617
DB 177 GlnGluTrpTrpGluLeuArgHisThrHisProValLeuTyrLeuPheTyrGluAspMet 196
QY 618 CATCGGACCTGGTGCAGTGGAGCAGTGGCCAGATTCCTGGGGGTGTCTCTGTGAC 677
DB 197 LysGluAsnProLysArgGluIleGlnLysIleLeuGluPheValGlyArgSerLeuPro 216
QY 678 AAGCCCGAGCTGGAAGCCCTGAGGACACTGC-----CACCACTGGTGGACAG 728
DB 217 GluGluThrValGluAspIleValGlnHisThrSerPheGlnGluMetLysAsnAsnAla 236
QY 729 TGCTGCAACGCTGAGGCCCTGCC-----GTG 755
DB 237 MetThrAsnTyrArgThrLeuProSerAspLeuAspHisSerIleSerAlaPheMet 256
QY 756 GGCCGGGGAAGAGTTGGCTGTGGAGGACATCTTACCCGCTCCATGAATGAAGTTT 815
DB 257 ArgLysGlyIleThrGlyAspTrpLysSerThrPheThrValAlaGlnAsnGluArgPhe 276
QY 816 GACTTGTGTATTAACAGAGATGGGAAAGTGTGACCTCAGCTTT 860
DB 277 GluAlaAspTyrAlaGluLysMetAlaGlyCysAsnLeuArgPhe 291

RESULT 28
STK2_RAT STANDARD; PRT; 296 AA.
AC Q9WUW9;
DT 01-DEC-2000, integrated into UniProtKB/Swiss-Prot.
DT 01-NOV-1999, sequence version 1.
DT 07-FEB-2006, entry version 29.
DE Sulfotransferase K2 (EC 2.8.2.-) (rSULT1C2A).
GN Name=Sultk2; Synonyms=Sultic2;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA].
RX TISSUE=Kidney;
RC MEDLINE=20329235; PubMed=10872834; DOI=10.1006/bbrc.2000.2744;
RA Xiangrong L., Joehnk C., Hartmann D., Scheestag F., Kroemer W.,
RA Gieselmann V.;
RT "Enzymatic properties, tissue-specific expression, and lysosomal
RT location of two highly homologous rat SULT1C2 sulfotransferases.";
RL Biochem. Biophys. Res. Commun. 272:242-250(2000).
CC -!- FUNCTION: Catalyzes the sulfate conjugation of many drugs,
CC xenobiotic compounds, hormones, and neurotransmitters.
CC -!- SUBCELLULAR LOCATION: Lysosome.
CC -!- TISSUE SPECIFICITY: Highly expressed in kidney and at lower levels
CC in stomach and liver. More specifically found in the epithelia of
CC proximal tubules of the kidney, of the bile duct, of the gastric
CC mucosa, and in hepatocytes.
CC -!- SIMILARITY: Belongs to the sulfotransferase family.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AJ238392; CAB41461.1; -; mRNA.
DR HSRP; P50224; 1CJM.
DR SMR; Q9WUW9; 12-296.
DR Ensembl; ENSRNOG0000031833; Rattus norvegicus.
DR RGD; 621064; Sult1c2.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase_1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
DR Lysosome; Transferase.
FT CHAIN 1 296 Sulfotransferase K2.
FT NP_BIND 49 54 /FtId=PRO_0000085136.
FT NP_BIND 131 139 PAPS (By similarity).
FT NP_BIND 194 230 PAPS (By similarity).

```



```
Db 277 AspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe 291
RESULT 31
Q2NL71 HUMAN PRELIMINARY; PRT; 295 AA.
AC Q2NL71
DT 07-FEB-2006, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, sequence version 1.
DE SULT1A1 protein.
GN Name=SULT1A1;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN NUCLEOTIDE SEQUENCE.
[1]
TX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Soufard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[2]
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Colon;
RC NIH MGC Project;
RG Submitted (DEC-2005) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BC110887; AAI10888.1; -; mRNA.
SQ SEQUENCE 295 AA; 34197 MW; 60E9D5BBA9159176 CRC64;

Alignment Scores:
Pred. No.: 4,42e-23 Length: 295
Score: 451.00 Matches: 93
Percent Similarity: 54.5% Conservative: 57
Best Local Similarity: 33.8% Mismatches: 107
Query Match: 10.2% Indels: 18
DB: 2 Gaps: 3

US-10-768-158-1 (1-2419) x Q2NL71_HUMAN (1-295)
QY 90 GCGTGGCGGCTCCCGCCCTTCGCGCGGGAGAGGAGGATCGCCACCTCCCGGTG 149
Db 17 GlyValProLeuLeuLysTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla 36
QY 150 CGGCCACGACGCTGTGATGCTACCTACCCCAAGTCCGCGCACGCTTGTCTGCAGGAG 209
Db 37 ArgProAspAspLeuLeuLeuSerThrTyrProLysSerGlyThrThrTrpValSerGln 56
QY 210 GTGGTCTACTTGGTGAGCCAGGCGCTGACCCCGATGATGCGGCTTGATGAACATCGAC 269
Db 57 IleLeuAspMetIleTyrGlnGlyAspLeuGluLysCysHisArgAlaProIlePhe 76

270 GAGCAGCTCCGGTCTCTGGAGTACCCACACGCG-----GGCTGGACATCATC 317
Db 77 MetArgValProPheLeuGluPheLysAlaProGlyIleProSerGlyMetGluThrLeu 96
QY 318 AAGAACTGACCTCTCTCCCGCTCATCAAGAGCCACCTGCGCTACCGCTTCTCCCTCT 377
Db 97 LysAspThrProAlaProArgLeuLeuLysThrHisLeuProLeuAlaLeuLeuProGln 116
QY 378 GACCTCCACAATGGAGACTCCAAGTCACTATATAGCTCGCAACCCCAAGGATCTCTGTG 437
Db 117 ThrLeuLeuAspGlnLysValLysValValTyrValAlaArgAsnAlaLysAspValAla 136
QY 438 GTGCTTATATATGTTCCACCGCTCTCTCGGGACCATGAGCTACCCAGGACCTTTCAA 497
Db 137 ValSerTyrTyrHisPheTyrHisMetAlaLysValHisProGluProGlyThrTrpAsp 156
QY 498 GAATTCTCGCGGAGTTTATGAATGATAAGCTGGGTACGGCTCCTGCTTTGAGCAGGTG 557
Db 157 SerPheLeuGluLysPheMetValGlyGluValSerTyrGlySerTyrTyrGlnHisVal 176
QY 558 CAGGAGTCTCGGAGACACCGCATGAGACTCGAACTGCTTTTCTCAAGTATGAAGACATG 617
Db 177 GlnGluTrpTrpGluLeuSerArgThrHisProValLeuTyrLeuPheTyrGluAspMet 196
QY 618 CATCGGACCTGGTGACGATGGTGGACAGCTGCGCCAGATTCCTGGGGGTGCTCTGTGAC 677
Db 197 LysGluAsnProLysArgGluIleGlnLeuGluPheValGlyArgSerLeuPro 216
QY 678 AAGCCCCAGCTGGAAGCCCTGACGAGCACTGCG-----CACAGCTGGTGGACACG 728
Db 217 GluGluThrValAspPheMetValGlnHisThrSerPheLysGluMetLysLysAsnPro 236
QY 729 TGCTGCAACGCTGAGGCCCTCGCC-----GTG 755
Db 237 MetThrAsnTyrThrThrValProGlnGluPheMetAspHisSerIleSerProPheMet 256
QY 756 GCGCGGGAAGAGTGGCTGTGGAAGGACATCTTCACCGCTTCCATGATGAGAAGTTT 815
Db 257 ArgLysGlyMetAlaGlyAspTrpLysThrThrPheThrValAlaGlnAsnGluArgPhe 276
QY 816 GACTTGGTGTATAAACAGAAGATGGAAAGTGTGACCTCACCTGTTT 860
Db 277 AspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe 291

RESULT 32
Q3ZC30 BOVIN PRELIMINARY; PRT; 295 AA.
AC Q3ZC30
DT 27-SEP-2005, integrated into UniProtKB/TrEMBL.
DT 27-SEP-2005, sequence version 1.
DT 07-MAR-2006, entry version 5.
DE Sulfotransferase, estrogen-prefering.
GN Name=STE;
OS Bos taurus (Bovine);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=Hereford; TISSUE=Reticulum;
RA Moore S., Alexander L., Brownstein M., Guan L., Lobo S., Meng Y.,
RA Taniguchi M., Wang Z., Yu J., Prange C., Schreiber K., Shermen C.,
RA Wagner L., Bala M., Barbazuk S., Barber S., Babakaiff R., Beland J.,
RA Chun E., Del Rio L., Gibson S., Hanson R., Kirkpatrick R., Liu J.,
RA Matsuo C., Mayo M., Santos R.R., Stott J., Tsai M., Wong D.,
RA Siddiqui A., Holt R., Jones S.J., Marra M.A.;
RL Submitted (AUG-2005) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
```

Db 260 AspValGlyAspTrpLysAsnHisPheThrValAlaLeuAsnGluLysPheAspMetHis 279

Qy 825 TATAAACAAGATGAGGAAAGTGTGACCTCAGCTTT 860
 |||::|||::|||
 Db 280 TyrGluGlnMetLysGlySerThrLeuLysPhe 291
 |||::|||::|||

RESULT 33

STIAL CANFA STANDARD; PRT; 295 AA.

AC Q29476;

DT 26-APR-2005, integrated into UniProtKB/Swiss-Prot.

DT 01-NOV-1996, sequence version 1.

DT 07-FEB-2006, entry version 29.

DE Sulfotransferase 1A1 (BC 2.8.2.1) (Aryl sulfotransferase) (Phenol
 sulfotransferase) (Phenol-sulfating phenol sulfotransferase) (P-PST).

GN Name=SULT1A1;

OS Canis familiaris (Dog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
 OC Canis.

NCBI_TaxID=9615;

[1]

RP NUCLEOTIDE SEQUENCE [MRNA].

RC STRAIN=Beagle; TISSUE=Liver;

RA Satsukawa M., Ogura K., Nakamura T., Watabe T.;

RT "Molecular cloning and sequencing of a dog liver cDNA (dPST-1)
 encoding a phenol sulfotransferase.";

RL Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.

[2]

RP NUCLEOTIDE SEQUENCE [MRNA], CHARACTERIZATION, AND TISSUE SPECIFICITY.

RC TISSUE=Liver;

CC MEDLINE=22049553; PubMed=12054462; DOI=10.1016/S0003-9861(02)00021-8;
 CC Tsai C., Morgenstern R., Swedmark S.;

RA "Canine sulfotransferase SULT1A1: molecular cloning, expression, and
 characterization.";

RL Arch. Biochem. Biophys. 401:125-133(2002).

CC -!- FUNCTION: Catalyzes the sulfate conjugation of catecholamines and
 phenolic drugs and neurotransmitters. Acts on p-nitrophenol (PNP),
 alpha-naphthol, dopamine, minoxidil and beta-estradiol but not
 dehydroepiandrosterone.

CC -!- CATALYTIC ACTIVITY: 3'-phosphoadenylyl sulfate + a phenol =
 CC adenosine 3',5'-bisphosphate + an aryl sulfate.

CC -!- SUBUNIT: Homodimer (By similarity).

CC -!- SUBCELLULAR LOCATION: Cytoplasm (By similarity).

CC -!- TISSUE SPECIFICITY: Ubiquitously expressed in canine tissues with
 CC highest expression in male and female liver.

CC -!- SIMILARITY: Belongs to the sulfotransferase family.

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License

EMBL; D29807; BAA06190.1; -; mRNA.

DR EMBL; AY069922; AAL57717.1; -; mRNA.

DR HSSP; P50224; 1CJM.

DR SMR; Q29476; 8-295.

DR Ensembl; ENSCAFG0000017122; Canis familiaris.

DR InterPro; IPR000863; Sulfotransferase.

DR Pfam; PF00685; Sulfotransferase 1.

DR ProDom; PD001218; Sulfotransferase; 1.

KW Catecholamine metabolism; Lipid metabolism; Steroid metabolism;
 KW Transferase.

CHAIN 1 295 Sulfotransferase 1A1.
 /FTid=PRO_0000085126.

FT NP_BIND 48 53 PAPS (By similarity).

FT NP_BIND 130 138 PAPS (By similarity).

FT NP_BIND 193 229 PAPS (By similarity).

FT NP_BIND 257 259 PAPS (By similarity).

FT ACT_SITE 108 108 Proton acceptor (By similarity).

SQ SEQUENCE 295 AA; 34115 MW; C297A9211A5609B6 CRC64;

Alignment Scores:
 Pred. No.: 5.2e-23 Length: 295
 Score: 450.00 Matches: 95


```
Qy 558 CAGGAGTCTCGGAGACCCGATGACGTCGACGCTGCTTTTCTCAAGTATGAAGACATG 617
Db 177 ArgglutrpTggluLeuSerArgThrHisProValleuTyLeuPhePheGluAspMet 196
Qy 618 CATCGGACCTGGTGTGACGATGCTGAGCAGCTGCCAGATTCCTCGGGGCTGCTGTGAC 677
Db 197 LysGluAsnProLysArgGluIleLysLysIleLeuGluPheLeuGlyArgSerLeuPro 216
Qy 678 AAGCCCGCAGCTGGAAGCCCTGACGAGCCTGCCACCGACGCTGGTGGACCCAG----- 728
Db 217 GluGluThrValAspArgIleAla--HisCysThrSerPheLysGluMetLysLysAsn 235
Qy 729 ---TGCTGCAACGCTGAGCCCTGCCCC----- 752
Db 236 ProMetThrAsnTySerThrIleProGluAsnIleMetAspHisAsnValSerProPhe 255
Qy 753 GTGGCGCGGGAAGAGTGGCTGTGAAGACATCTTCACGCTCCATCAATGAGAAG 812
Db 256 MetArgargGlyValAlaGlyAspTrpLysThrThrPheThrValAlaGlnHisGluTyr 275
Qy 813 TTTCAGCTTGGTGTATAAACAGAGATGGGAAGTGTGACCTCACGTTT 860
Db 276 PheGluAlaAspTyrAlaGluLysMetAlaGlyCysGluLeuThrPhe 291

RESULT 35
ID ST1C1 HUMAN STANDARD; PRT; 296 AA.
AC 000338;
DT 01-DEC-2000, integrated into UniProtKB/Swiss-Prot.
DT 01-JUL-1997, sequence version 1.
DT 07-MAR-2006, entry version 52.
DE Sulfotransferase 1C1 (EC 2.8.2.-) (SULT1C#1) (SULT1C2) (humsULTC2).
GN Name=SULT1C1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RN NUCLEOTIDE SEQUENCE.
RC TISSUE=Fetal liver, and Fetal spleen;
RX MEDLINE=97312707; PubMed=9169148; DOI=10.1006/geno.1997.4683;
RA Her C., Kaur G.P., Athwal R.S., Weinshilboum R.M.;
RT "Human sulfotransferase SULT1C1: cDNA cloning, tissue-specific
expression, and chromosomal localization.";
RL Genomics 41:467-470(1997).
RN [2]
RN NUCLEOTIDE SEQUENCE.
RC TISSUE=Fetal liver;
RX MEDLINE=98297811; PubMed=9635888; DOI=10.1093/carcin/19.5.951;
RA Yoshinari K., Nagata K., Shimada M., Yamazoe Y.;
RT "Molecular characterization of SULT1C1-related human sulfotransferase.";
RL Carcinogenesis 19:951-953(1998).
RN [3]
RN NUCLEOTIDE SEQUENCE, AND CHARACTERIZATION.
RC TISSUE=Fetal lung;
RX MEDLINE=99069375; PubMed=9852044; DOI=10.1074/jbc.273.51.33929;
RA Sakakibara Y., Yanagisawa K., Katafuchi J., Ringer D.P., Takami Y.,
RA Nakayama T., Suiko M., Liu M.-C.;
RT "Molecular cloning, expression, and characterization of novel human
SULT1C sulfotransferases that catalyze the sulfonation of N-hydroxy-2-
acetylaminofluorene.";
RL J. Biol. Chem. 273:33929-33935(1998).
RN [4]
RN NUCLEOTIDE SEQUENCE, AND CHARACTERIZATION.
RC TISSUE=Stomach;
RX MEDLINE=99410886; PubMed=10481272; DOI=10.1016/S1357-2725(99)00038-2;
RA Hehonah N., Zhu X., Brix L., Bolton-Grob R., Barnett A., Windmill K.,
RA McManus M.;
RT "Molecular cloning, expression, localisation and functional
characterisation of a rabbit SULT1C2 sulfotransferase.";
RL Int. J. Biochem. Cell Biol. 31:869-882(1999).
RN [5]
```

```
RP NUCLEOTIDE SEQUENCE, AND ALTERNATIVE SPLICING.
RX MEDLINE=20247255; PubMed=10783263; DOI=10.1006/geno.2000.6150;
RA Freimuth R.R., Rafiloglanis R.B., Wood T.C., Moon E., Kim U.-J., Xu J.,
RA Sciliano M.J., Weinshilboum R.M.;
RT "Human sulfotransferases SULT1C1 and SULT1C2: cDNA characterization,
gene cloning, and chromosomal localization.";
RL Genomics 65:157-165(2000).
RN [6]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM SHORT).
RA Kalline N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S.,
RA Koundinya M., Raphael J., Moreira D., Kelley T., Labaer J., Lin Y.,
RA Phelan M., Farmer A.;
RT "Cloning of human full-length cDNAs in BD Creator(TM) system donor
vector.";
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
RN [7]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
TS TISSUE=Kidney;
RX MEDLINE=22398257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Helton E., Kettelman M., Maman A.C., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Catalyzes the sulfate conjugation of many drugs,
xenobiotic compounds, hormones, and neurotransmitters. May be
involved in the activation of carcinogenic hydroxylamines. Shows
activity towards p-nitrophenol and N-hydroxy-2-acetylmino-
fluorene (N-OH-2AFL).
CC -!- SUBCELLULAR LOCATION: Cytoplasm.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Short;
CC IsoId=O000338-1; Sequence=Displayed;
CC Names=Long;
CC IsoId=O000338-2; Sequence=VSP_006303;
CC -!- TISSUE SPECIFICITY: Found in adult stomach, kidney and thyroid
gland, and in fetal kidney and liver.
CC -!- SIMILARITY: Belongs to the sulfotransferase family.
CC
CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; U66036; AAC51285.1; -; mRNA.
DR EMBL; AB008164; BA28346.1; -; mRNA.
DR EMBL; AF026303; AAC00409.1; -; mRNA.
DR EMBL; AF186251; AAF72799.1; -; mRNA.
DR EMBL; AF186252; AAF72800.1; -; mRNA.
DR EMBL; AF186253; AAF72801.1; -; mRNA.
DR EMBL; AF186254; AAF72802.1; -; mRNA.
DR EMBL; AF186255; AAF72803.1; -; mRNA.
DR EMBL; AF186256; AAF72804.1; -; mRNA.
DR EMBL; AF186262; AAF72805.1; -; Genomic DNA.
DR EMBL; AF186258; AAF72805.1; JOINED; Genomic DNA.
DR EMBL; AF186260; AAF72805.1; JOINED; Genomic DNA.
DR EMBL; AF186261; AAF72805.1; JOINED; Genomic DNA.
DR EMBL; AF186262; AAF72806.1; -; Genomic DNA.
DR EMBL; AF186258; AAF72806.1; JOINED; Genomic DNA.
```


DR EMBL; AF186259; AAF72806.1; JOINED; Genomic DNA.
 DR EMBL; AF186260; AAF72806.1; JOINED; Genomic DNA.
 DR EMBL; AF186261; AAF72806.1; JOINED; Genomic DNA.
 DR EMBL; BT006951; AAF35597.1; -; mRNA.
 DR EMBL; BC005353; AAH05353.1; -; mRNA.
 DR PDB; 2ETG; X-ray; A/B=1-296.
 DR Ensembl; ENSG00000198203; Homo sapiens.
 DR H-invDB; HIX0002353; -.
 DR HGNC; HGNC:11456; SULT1C1.
 DR MIM; 602385; gene.
 DR GO; GO:0005737; C:cytoplasm; TAS.
 DR GO; GO:0008146; F:sulfotransferase activity; TAS.
 DR GO; GO:0009308; P:amine metabolism; TAS.
 DR InterPro; IPR000863; Sulfotransferase.
 DR Pfam; PF00685; Sulfotransfer_1; 1.
 DR ProDom; PD001218; Sulfotransferase; 1.
 DR 3D-structure; Alternative splicing; Polymorphism; Transferase.
 CHAIN 1 296
 FT NP_BIND 49 54
 FT NP_BIND 131 139
 FT NP_BIND 194 230
 FT NP_BIND 258 260
 FT ACT_SITE 109 109
 FT VARSPIC 93 113
 FT
 FT VARIANT 128 128
 FT
 FT VARIANT 255 255
 FT
 FT STRAND 14 16
 FT TURN 17 18
 FT STRAND 19 21
 FT HELIX 23 27
 FT TURN 28 28
 FT HELIX 29 33
 FT TURN 34 34
 FT TURN 39 40
 FT STRAND 42 46
 FT TURN 48 49
 FT STRAND 50 51
 FT HELIX 52 64
 FT TURN 65 65
 FT STRAND 82 82
 FT TURN 83 84
 FT STRAND 88 89
 FT TURN 90 90
 FT HELIX 93 99
 FT STRAND 100 100
 FT STRAND 102 103
 FT STRAND 105 108
 FT TURN 112 114
 FT STRAND 115 115
 FT TURN 117 117
 FT HELIX 118 121
 FT TURN 122 123
 FT STRAND 125 130
 FT HELIX 133 146
 FT TURN 148 149
 FT STRAND 150 150
 FT STRAND 155 155
 FT HELIX 156 164
 FT TURN 166 166
 FT STRAND 168 169
 FT TURN 170 170
 FT STRAND 173 183
 FT TURN 184 186
 FT STRAND 187 193
 FT HELIX 194 199
 FT TURN 201 211
 FT HELIX 212 213
 FT TURN 218 218
 FT HELIX 219 227

FT TURN 228 228
 FT STRAND 229 229
 FT HELIX 230 235
 FT STRAND 262 262
 FT TURN 264 264
 FT HELIX 265 268
 FT STRAND 271 285

Alignment Scores: 6.13e-23 Length: 296
 Pred. No.: 449.00 Matches: 94
 Score: 54.6% Conservative: 61
 Percent Similarity: 33.1% Mismatches: 105
 Best Local Similarity: 10.2% Indels: 24
 Query Match: 1 Gaps: 4
 DB: 1

US-10-768-158-1 (1-2419) x STIC1_HUMAN (1-296)
 QY 72 AAGTACTTCGAGTTCCATGGCTGCGCTCCGCCCTTCTCGCGGGAAGATGGAGGAG 131
 DB 12 LysLeuLysGluValGluGlyThrLeuLeuGlnProAlaThrValAspAanTrpSerGln 31
 QY 132 ATCCCAACTTCCCGTGGCGCCAGCGACGTGTGATCGTCACCTACCCCAAGTCGCGC 191
 DB 32 IleGlnSerPheGluAlaLysProAspAspLeuLeuLeuCysThrTyProLysAlaGly 51
 QY 192 ACCAGCTTGTGGAGAGGTGCTACTTGTGAGCGCGCGCTGACCCCGATGAGATC 251
 DB 52 ThrTrpIleGlnGluIleValAspMetIleGluGlnAenGlyAspValGluLysCys 71
 QY 252 GGCTTGATGAACATCGACGAGCAGCTCCCGTCTCTGGAGTAC-----CCACAGCCG 302
 DB 72 GlnArgAlaIleIleGlnHisArgHisProPheIleGluTrpAlaArgProProGlnPro 91
 QY 303 ---GGCTGGACATCATCAAGAACTGACCTCTCCCGCTCATCAAGAGCCACTGCCCC 359
 DB 92 SerGlyValGluLysAlaLysAlaMetProSerProArgIleLeuLysThrHisLeuSer 111
 QY 360 TACCGCTTCTGCGCTGACCTCCACATGGAGATCCCAAGGTCTATCTATATGCTGCGC 419
 DB 112 ThrGlnLeuLeuProProSerPheTrpGluAsnAsnCysLysPheLeuTyValAlaArg 131
 QY 420 AACCCCAAGGATCTGGTGTCTTATTATTCAGTCTCCACCGCTCTCTGCGGACCATAGC 479
 DB 132 AsnAlaLysAspCysMetValSerTyThrHisPheGlnArgMetAsnHisMetLeuPro 151
 QY 480 TACCGAGCACCTTTCAAGAAATTCGCGGAGGTATTATGAATGATATAGTGGCTACGCG 539
 DB 152 AspProGlyThrTrpGluGluTyPheGluThrPheIleAsnGlyLysValValTrpGly 171
 QY 540 TCCTGGTTTGAGCAGCTGAGGAGTCTTGGGAGCACCGCATGGACTCGAACGTCCTTTT 599
 DB 172 SerTrpPheAspHisValLysGlyTrpTrpGluMetLysAspArgHisGlnIleLeuPhe 191
 QY 600 CTCAAGTATGAACATCGGACCTGGTGCAGCATGGTGGAGCAGCTGGCGGACGATTC 659
 DB 192 LeuPheTyrgluaspIleLysArgAspProLysHisGluIleArgLysValMetGlnPhe 211
 QY 660 CTGGGGGTGCTGTGTGACAAGCCCGAGCTGGAAGCCCTGACGCGAGCAC----- 707
 DB 212 MetGlyLysLysValAspGluThrValLeuAspLysIleValGlnGluThrSerPheGlu 231
 QY 708 -----TCCACACAGCTGGTGACGAG 728
 DB 232 LysMetLysGluAsnProMetThrAsnArgSerThrValSerLysSerIleLeuAspGln 251
 QY 729 TGCTGCAACGCTAGGCCCTCGCGCGGCGGAAGAGTTGGGCTGTGTGAAGAGCATC 788
 DB 252 SerIleSerSerPhe-----MetArgLysGlyThrValGlyAspTrpLysAsnHis 268
 QY 789 TTCACCGCTCTCCATGAATGAGAGTTTGACTGCTGTATTAACAGAGATGGGAAAGTGT 848
 DB 269 PheThrValAlaGlnAsnGluArgPheAspGluIleTyArgArgLysMetGluGlyThr 288


```

Qy 594 CTTTTTCTCAAGTATGAGACATGCATCGGACCTGGTGGAGCTGGCC 653
Db 196 LeuTyrLeuPheTyrGluAspMetLysAsnProLysHisGluLeuAla 215
Qy 654 AGATTCTCGGGGTCTCTGTGCAAGGCCAGCTGGAAGCCCTGACGGAGCACTGCCAC 713
Db 216 GluPheLeuGlyLysLysLeuAspLysValLeuAspLysIleValHisTyrThrSer 235
Qy 714 CAGCTGGTG-----GACCACTGTGCAACCTGAGGCCCTGCCCC----- 752
Db 236 PheAspValMetLysGlnAsnProMetAlaAsnTyrSerSerIleProAlaGluIleMet 255
Qy 753 -----GTGGCGGGGAGAGCTGGCTGGTGGAGGACATCTTC 791
Db 256 AspHisSerIleSerProPheMetArgLysGlyAlaValGlyAspTrpLysHisPhe 275
Qy 792 ACCGTCTCCATGAATGAGAGCTTTGACTGTGTATATAACAGAGATGGGAAAGTGTGAC 851
Db 276 ThrValAlaGlnAsnGluArgPheAspGluAspTyrLysLysLysMetThrAspThrArg 295
Qy 852 CTCACGTTGACTTT 866
Db 296 LeuThrPheHisPhe 300

RESULT 37
Q95MF8_PIG PRELIMINARY; PRT; 295 AA.
AC Q95MF8;
DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT 01-DEC-2001, sequence version 1.
DT 07-FEB-2006, entry version 14.
DE Estrogen sulfotransferase.
OS Sus scrofa (pig).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
OC Sus.
OX NCBI_TaxID=9823;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22345564; PubMed=12457956; DOI=10.1016/S0739-7240(02)00172-8;
RA Kim J.G., Vallet J.L., Rohrer G.A., Christenson R.K.;
RT "Characterization of porcine uterine estrogen sulfotransferase.";
RL Domest. Anim. Endocrinol. 23:493-506(2002).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
DR EMBL; AF389855; AAK72967.1; -, mRNA.
DR HSP; P49888; 1HY3.
DR SMR; Q95MF8; 4-293.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase 1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
DR TransFam;
SQ SEQUENCE 295 AA; 34940 MW; 7677C28A91B38167 CRC64;

Alignment Scores:
Pred. No.: 8.49e-23 Length: 295
Score: 447.00 Matches: 96
Percent Similarity: 54.3% Conservative: 56
Best Local Similarity: 34.3% Mismatches: 106
Query Match: 10.2% Indels: 22
DB: 2 Gaps: 4

US-10-768-158-1 (1-2419) x Q95MF8_PIG (1-295)
Qy 66 GAGAGCAAGTACTTCGAGTTCT-----CATGGCTGGCGGCTGGCCCTTCTGCG 113
Db 5 LysSerAlaTyrLeuAspTyrPheGlyArgIleHisGlyIleLeuLeuTyrLysLysPhe 24

```

```

Qy 114 CGCGGGAAGATGGAGGAGATCGCCAACTTCCCGTGGCGCCAGCGAGCTGTGGATCGTC 173
Db 25 IleGluTyrTrpAsnAspValGluThrPheGluAlaArgProAspLeuValIleVal 44
Qy 174 ACCTACCCCAAGTCGGCACCAGCTTCTGCTCAGAGAGTGGTCTACTTGTGGAGCCAGGC 233
Db 45 ThrTyrProLysSerGlyThrTrpValSerGluIleValTyrMetIleTyrThrGlu 64
Qy 234 GCTGACCCCATGAGATCGCTTGTATGAACATGACGAGCAGCTCCCGCTCTG----- 287
Db 65 GlyAspValGluLysCysLysGluAspThrIlePheAsnArgIleProTyrLeuGluCys 84
Qy 288 -----GAGTACCCACAGCGCGGCTGGACATCATCAAGGAAGTCTCCCGCGCTC 341
Db 85 ArgThrGluAsnValMetAsnGlyValLysGlnLeuLysGlnMetAlaSerProArgIle 104
Qy 342 ATCAAGAGCCACCTGCCCTACCGCTTCTCCCTCTGACCTCCACATGAGAGCTCCCAAG 401
Db 105 ValLysSerHisLeuProProGluLeuLeuProValSerPheTrpGluLysAsnCysLys 124
Qy 402 GTCATCTATATGGCTGCCAACCCCAAGGATCTGGTGTCTGTTATATCATAGTTCCACCGC 461
Db 125 IleIleTyrValCysArgAsnAlaLysAspValValSerTyrTyrPhePheLeu 144
Qy 462 TCTCTCGGACCATGAGTACCGGACCTTTCAAGAATCTCTCGGAGGTTTATGAAT 521
Db 145 MetValThrAlaAsnProAspProGlySerPheGlnAspPheValGluLysPheMetAsp 164
Qy 522 GATAAGCTGGCTACGGCTCTCTGTTGACAGCTGCAGGAGTTCTGGGAGCACCGCATG 581
Db 165 GlyGluValProTyrGlySerTyrTyrLysHisThrLysSerTrpTrpGluLysArgThr 184
Qy 582 GACTCGAACCTGCTTTTCTCAAGTATGAAGATGATCATCGGACCTGGTGTGAGATGGTG 641
Db 185 AsnProGlnValLeuPheIlePheTyrGluAspMetLysGluAsnIleArgLysGluVal 204
Qy 642 GAGCAGCTGCCAGATTCTCTGGGGGTCTCTGTGACAAGCCCGAGCTGGAGCCCTGACG 701
Db 205 MetArgLeuIleGluPheLeuGlyArgLysAlaSerAspGluLeuValAspLysIle 224
Qy 702 GAGCAGCTGCCAC-----CAGCTGGTG 722
Db 225 LysHisThrSerPheGlnGluMetLysAsnAsnProSerThrAsnTyrThrThrLeuPro 244
Qy 723 GACCACTGCTGCAACGCTGAGCGCTGCGCC-----GTGGCGCGGAGAGATGGCTGTGG 779
Db 245 AspGluValMetAsnGlnLysValSerAlaPheMetArgLysGlyIleAlaGlyAspTrp 264
Qy 780 AAGGACATCTTCCCGTCTCCATGAATGAGAAGTTTGACTTGGTGTATAACAGAACATG 839
Db 265 LysAsnTyrPheThrValAlaLeuAsnGluLysPheAspIleHisTyrGluGlnMet 284

RESULT 38
ST1B1_RAT
ID ST1B1_RAT STANDARD; PRT; 299 AA.
AC P52847;
DT 01-OCT-1996, integrated into UniProtKB/Swiss-Prot.
DT 15-DEC-1998, sequence version 2.
DT 07-FEB-2006, entry version 42.
DE Sulfotransferase family cytosolic 1B member 1 (EC 2.8.2.-)
DE (Sulfotransferase 1B) (DOPA/tyrosine sulfotransferase).
GN Names:St1b1; Synonyms=St1b1;
OS Rattus norvegicus (Rat);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA], AND PROTEIN SEQUENCE OF 108-117; 126-134
RP AND 260-266.
RC STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=96107201; PubMed=8530477; DOI=10.1074/jbc.270.51.30470;
RA Sakakibara Y., Takami Y., Zwieb C., Nakayama T., Suiko M.,

```



```

CC CC -!- CATALYTIC ACTIVITY: 3'-phosphoadenylyl sulfate + estrone =
CC CC adenosine 3',5'-bisphosphate + estrone 3-sulfate.
CC CC -!- SUBUNIT: Homodimer.
CC CC -!- SUBCELLULAR LOCATION: Cytoplasm.
CC CC -!- INDUCTION: By progesterone.
CC CC -!- SIMILARITY: Belongs to the sulfotransferase family.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; M54942; AAA30679.1; -; mRNA.
CC DR EMBL; X56395; CAA39806.1; -; mRNA.
CC DR PIR; S29045; S29045.
CC DR HSSP; P49888; 1HV3.
CC DR SMR; P19217; 4-293.
CC DR InterPro; IPR000863; Sulfotransferase.
CC DR Pfam; PF00685; Sulfotransferase; 1.
CC DR ProDom; PD001218; Sulfotransferase; 1.
CC KW Direct protein sequencing; Lipid-binding; Phosphorylation;
KW Steroid-binding; Transferase.
CC CHAIN 1 295 Estrogen sulfotransferase.
CC FT NP_BIND 48 53 PAPS (By similarity).
CC FT NP_BIND 130 138 PAPS (By similarity).
CC FT NP_BIND 193 229 PAPS (By similarity).
CC FT NP_BIND 257 259 PAPS (By similarity).
CC FT ACT_SITE 108 108 Proton acceptor (By similarity).
CC FT MOD_RES 216 216 Phosphoserine (by PKA) (Potential).
CC FT MOD_RES 228 228 Phosphoserine (by PKA) (Potential).
CC FT CONFLICT 117 117 S -> E (in Ref. 2).
CC FT CONFLICT 171 171 S -> E (in Ref. 2).
CC FT CONFLICT 248 248 M -> Q (in Ref. 2).
CC SQ SEQUENCE 295 AA; 34640 MW; 69147C73146913FD CRC64;

Alignment Scores:
Pred. No.: 1e-22 Length: 295
Score: 446.00 Matches: 100
Percent Similarity: 53.8% Conservative: 57
Best Local Similarity: 34.2% Mismatches: 111
Query Match: 10.1% Indels: 24
DB: 1 Gaps: 5

US-10-768-158-1 (1-2419) x STIE1_BOVIN (1-295)
Qy 48 AGCACCCCGGGAGTTCAGAGCAAGTACTTCAGATTCCATGGCTGCGCTGCCGCC 107
Db 3 SerSerLysProSerPheSerAspTy:RPhelyLysLeuGlyGlyLe------Pro 19
Qy 108 TTCTGCGCGGGAAGATGGAGGAGATGCCAAC-----TTCCCGGTGCGGCCAGC 158
Db 20 MetTyLysLysPheIleGluGlnPheHisAsnValGluGluPheGluAlaArgProAsp 39
Qy 159 GACGTGTGGATCGTCACCTACCCAGTCCCGGACCCAGCTTCGTCGAGGAGTGTCTAC 218
Db 40 AspLeuValIleValThrTyProLysSerGlyThrTrpLeuSerGluIleCys 59
Qy 219 TTGTGTGACGACGCGCTGACCCGATGAGATCGCTTGATGAACATCGACGAGCTC 278
Db 60 MetIleTyAsnAsnGlyAspValGluLysCysLysGluAspValIlePheAsnArgVal 79
Qy 279 CCGTCTCTG-----GAGTACCCACAGCGCGCTGGACATCATCAAGAACTG 326
Db 80 ProTyLeuGluCysSerThrGluHisValMetLysGlyValLysGlnLeuAsnGluMet 99
Qy 327 ACCTCTCCCGCTCATCAAGACCATGCGCTTTCGCCCTTCGACCTCCAC 386
Db 100 AlaSerProArgIleValLysSerHisLeuProValLysLeuLeuProValSerPheTrp 119
Qy 387 AATGAGAGCTCAAGGTTCATCTATATGGCTCGCAACCCCAAGGATCTGTGTCTTAT 446
Db 120 GlulysAsnCysLysIleIleTyLeuSerArgAsnAlaLysAspValValSerTy 139
Qy 447 TATCAGTTCCACCGCTCTCTCGCGGACCATGAGCTACCGAGGCACCTTTCAGAATTCTGC 506

```

```

Db 140 TyrPheLeuIleLeuMetValThrAlaIleProAspProAspSerPheGlnAepPheVal 159
Qy 507 CGGAGGTTTATGAATATAAGCTGGCTACGGCTCCTGGTTTGAGCACGTCGACGAGTTC 566
Db 160 GlulysPheMetAspGlyGluValProTyGlySerTrpPheGluHisThrLysSerTrp 179
Qy 567 TGGGAGCACCCGATGGACTCGAACGTCCTTTTCTCAAGTATGAAGACATGCATCGGGAC 626
Db 180 TrpGluLysSerLysAsnProGlnValLeuPheLeuPheTyGluAspMetLysGluAsn 199
Qy 627 CTGGTGACGATGGTGGAGCAGCTGGCCAGATTCTCTGGGGGTGCTCTGTGACAAAGCCGAC 686
Db 200 IleArgLysGluValMetLysLeuLeuGluPheLeuGlyArgLysAlaSerAspGluLeu 219
Qy 687 CTGGAAGCCCTGACGGACGATGCCAC----- 713
Db 220 ValAspLysIleIleLysHisThrSerPheGlnGluMetLysAsnAsnProSerThrAsn 239
Qy 714 -----CAGCTGTGGTGGACAGTGTGCAACGCTGAGGCCCTGCC---GTGGGCGCGGGA 764
Db 240 TyrThrThrLeuProAspGluValMetAsnGlnLysValSerProPheMetArgLysGly 259
Qy 765 AGAGTTGGCTGTGGAGGACATCTTCACCGTCTCCATGATGAAGAGTTTGACTTGGTG 824
Db 260 AspValGlyAspTrpLysAsnHisPheThrValAlaLeuAsnGluLysPheAspMetHis 279
Qy 825 TATAAACAGAAAGATGGGAAAGTGTGACCTCACGCTT 860
Db 280 TyrGluGlnGlnMetLysGlySerThrLeuLysPhe 291

RESULT 41
STIE1_HUMAN
ID STIE1_HUMAN STANDARD; PRT; 296 AA.
AC Q43704; O15497; Q95F11; Q9UK34;
DT 04-JAN-2005, integrated into UniProtKB/Swiss-Prot.
DT 07-FEB-2006, sequence version 2.
DE Sulfotransferase family cytosolic 1B member 1 (EC 2.8.2.-)
DE (Sulfotransferase 1B2) (Thyroid hormone sulfotransferase).
GN Name=STIE1B1; Synonyms=STIEB2, SULT1B2;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE, AND FUNCTION.
RC TISSUE=Liver;
RX MEDLINE=98104061; PubMed=9443824;
RA Fujita K., Nagata K., Ozawa S., Sasano H., Yamazoe Y.;
RT "Molecular cloning and characterization of rat STIE1 and human STIE2
cDNAs, encoding thyroid hormone sulfotransferases.";
RL J. Biochem. 122:1052-1061(1997).
RN [2]
RP NUCLEOTIDE SEQUENCE, FUNCTION, AND TISSUE SPECIFICITY.
RC TISSUE=Liver;
RX MEDLINE=98130699; PubMed=9463486;
RA Wang J., Palany J.L., Palany C.N.;
RT "Expression and characterization of a novel thyroid hormone-sulfating
form of cytosolic sulfotransferase from human liver.";
RL Mol. Pharmacol. 53:274-282(1998).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Bone marrow;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,

```

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Beak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzyzanski M.I., Skalska U., Smalus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4].
RP NUCLEOTIDE SEQUENCE OF 1-73.
RA Falany C.N., Wang J.;
RT "Mapping of the SULT1B2 gene to human chromosome 4q11-13.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Catalyzes the sulfate conjugation of many hormones,
CC neurotransmitters, drugs and xenobiotic compounds. Sulfonation
CC increases the water solubility of most compounds, and therefore
CC their renal excretion, but it can also result in bioactivation to
CC form active metabolites. Sulfates dopamine, small phenols such as
CC 1-naphthol and p-nitrophenol and thyroid hormones, including 3,3'-
CC diiodothyronine, triiodothyronine, reverse triiodothyronine and
CC thyroxine.
CC -1- SUBUNIT: Monomer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasm (By similarity).
CC -1- TISSUE SPECIFICITY: Highly expressed in the liver, peripheral
CC blood leukocytes, colon (mucosal lining), small intestine
CC (jejunum) and spleen. A lesser expression was observed in the
CC lung, placenta and thymus.
CC -1- SIMILARITY: Belongs to the sulfotransferase family.
CC
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC -----
DR EMBL: D89479; BAA24547.1; -; mRNA.
DR EMBL: U95726; AAB65154.1; -; mRNA.
DR EMBL: BC010895; AAH10895.1; -; mRNA.
DR EMBL: AF184894; AAF05917.1; -; Genomic_DNA.
DR PIR: JC5885; JCS5885.
DR PDB: 1XV1; X-ray; A/B=1-296.
DR Ensembl: ENSG00000173597; Homo sapiens.
DR HGNC: HGNC:17845; SULT1B1.
DR MIM: 608436; Gene.
DR GO: GO:0005229; Cytoplasm; NAS.
DR GO: GO:0008146; F:sulfotransferase activity; IDA.
DR GO: GO:0018958; P:phenol metabolism; IDA.
DR GO: GO:0042403; P:thyroid hormone metabolism; IDA.
DR InterPro: IPR000863; Sulfotransferase.
DR Pfam: PF00685; Sulfotransferase_1; 1.
DR ProDom: PD001218; Sulfotransferase; 1.
KW 3D-structure; Lipid metabolism; Polymorphism; Steroid metabolism;
KW Transferase.
FT CHAIN 1 296 Sulfotransferase family cytosolic 1B
FT member 1.
FT /FTID=PRO_0000085161.
FT NP_BIND 48 53 PAPS (By similarity).
FT NP_BIND 131 139 PAPS (By similarity).
FT NP_BIND 194 230 PAPS (By similarity).
FT ACT_SITE 109 109 Proton acceptor (By similarity).
FT BINDING 259 259 PAPS (By similarity).
FT VARIANT 145 145 L -> V (in dbSNP:11569736).
FT /FTID=VAR_020886.
FT CONFLICT 183 183 K -> R (in Ref. 3).
FT CONFLICT 186 186 E -> G (in Ref. 1).
FT TURN 5 9
FT STRAND 13 15
FT TURN 16 17
FT STRAND 18 20
FT STRAND 22 24
FT HELIX 25 26

FT HELIX 28 32
FT TURN 33 33
FT TURN 38 39
FT STRAND 41 46
FT TURN 47 48
FT STRAND 49 50
FT TURN 51 62
FT HELIX 63 65
FT TURN 67 68
FT HELIX 69 71
FT STRAND 72 73
FT HELIX 75 78
FT STRAND 79 79
FT STRAND 81 81
FT TURN 82 83
FT STRAND 86 86
FT TURN 87 89
FT STRAND 90 90
FT HELIX 93 97
FT TURN 98 99
FT STRAND 100 100
FT STRAND 102 103
FT STRAND 105 108
FT TURN 112 114
FT STRAND 115 115
FT HELIX 117 121
FT TURN 122 123
FT STRAND 125 130
FT HELIX 133 146
FT STRAND 147 147
FT TURN 148 149
FT STRAND 150 150
FT HELIX 156 164
FT TURN 165 166
FT STRAND 168 168
FT TURN 169 170
FT HELIX 173 185
FT TURN 186 186
FT STRAND 187 187
FT STRAND 189 193
FT HELIX 194 199
FT HELIX 201 211
FT TURN 212 213
FT HELIX 218 227
FT TURN 228 228
FT STRAND 229 229
FT HELIX 230 235
FT TURN 237 239
FT STRAND 240 241
FT TURN 242 243
FT STRAND 244 244
FT TURN 246 248
FT STRAND 249 249
FT TURN 251 253
FT STRAND 254 254
FT STRAND 256 257
FT STRAND 262 263
FT HELIX 264 267
FT TURN 268 268
FT HELIX 271 285
FT TURN 286 287
FT STRAND 288 289
FT STRAND 294 294
SQ SEQUENCE 296 AA; 34899 MW; AFE61B21DBD782C CRC64;

Alignment Scores:

Pred. No.: 1,51e-22
Score: 443.50
Percent Similarity: 53.8%
Best Local Similarity: 33.9%
Query Match: 10.1%
DB: 1

Length: 296
Matches: 94
Conservative: 55
Mismatch: 109
Indels: 19
Gaps: 3

US-10-768-158-1 (1-2419) x ST1B1_HUMAN (1-296)	
QY	87 CATGGCGTGGCGTCCCGCCCTTCTGCGCGGGAGATGGAGGAGATCGCAACTTCCG 146
Db	: : : : :
Db	16 HisGlyTyrProMetThrCysAlaPheAlaSerAenTrpGluLysIleGluGlnPheHis 35
QY	147 GTGGCGCCAGCGAGTGTGGATCGTACCTACCCAGTCCGCGCAGCAGCTTCTGCAG 206
Db	: : : : :
Db	36 SerArgProAspApeIleValIleAlaThrTyrProLysSerGlyThrTrpValSer 55
QY	207 GAGTGTGTCTACTTGGTCAGCGCGCTGACCCCGATGATCGGCTTGATGAACAT 266
Db	: : : : :
Db	56 GluIleIleApeMetIleLeuAsnAspGlyAspIleGluLysCysLysArgGlyPheIle 75
QY	267 GACGAGAGCTCCCGTCTCTGAGPACCCACAGCGCGCGCTG-----GAC 311
Db	: : : : :
Db	76 ThrGluLysValProMetLeuGluMetThrLeuProGlyLeuArgThrSerGlyIleGlu 95
QY	312 ATCATCAGGAAGTACCTCTCCCGCTCATCAGAGCCACCTACCGCTTCTCTG 371
Db	: : : : :
Db	96 GlnLeuGluLysAsnProSerProArgIleValLysThrHisLeuProThrAspLeuLeu 115
QY	372 CCCTCTGACCTCCCAATGGAGATCCCAAGTCTATATGGTCTGCGCAACCCCAAGGAT 431
Db	: : : : :
Db	116 ProlLysSerPheTrpGluAsnAsnCysLysMetIleTyrLeuAlaAaGAsnAlaLysAsp 135
QY	432 CTGTGGTGTCTTATTATACGTTCCACCGCTCTCTGCGGACCATGAGCTACCGGACCC 491
Db	: : : : :
Db	136 ValSerValSerTyrTyrHisPheAspLeuMetAenAenLeuGlnProPheProGlyThr 155
QY	492 TTTCAGAAATTCGCGGAGGTTTATGAATATGAATGAAGCTGGCTACGGCTCTGTTGAG 551
Db	: : : : :
Db	156 TrpGluGluTyrLeuGluLysPheLeuThrGlyLysValAlaTyrGlySerTrpPheThr 175
QY	552 CACGTGCGAGGAGTCTCTGGGACCGCATGCTGCAACGCTTTTCTCAAGTATGAA 611
Db	: : : : :
Db	176 HisValLysAenTrpTrpLysLysGluGluHisProlIleLeuPheLeuTyrTrpGlu 195
QY	612 GACATGTCATCGGACCTGGTGCAGTGTGGAGCAGCTGGCAGATCTCTGGGGTGTCC 671
Db	: : : : :
Db	196 AspMetLysGluAenProLysGluGluIleLysLysIleIleArgPheLeuGluLysAen 215
QY	672 TGTGACAGGCCAGCTGAGACCTGACGAGCACTGCCACCGCTGGTG----- 722
Db	: : : : :
Db	216 LeuAsnAspGluIleLeuAspArgIleIleHisThrSerPheGluValMetLysAsp 235
QY	723 GACCAGTGTCTGCAAGCTGAGCGCTGCC----- 752
Db	: : : : :
Db	236 AsnProLeuValAenTyrThrHisLeuProThrThrValMetAspHisSerLysSerPro 255
QY	753 ---GTGGCGCGGGAAGATTGGGTGTGGAGGACATCTTACCGTCTCCATGAATGAG 809
Db	: : : : :
Db	256 PheMetArgLysGlyThrAlaGlyAspTrpLysAenTyrPheThrValAlaGlnAsnGlu 275
QY	810 AAGTTTGACTGTGTGATAACACAGATGGGAAGTGTGACCTTCACGTTT 860
Db	: : : : :
Db	276 LysPheAspAlaIleTyrGluThrGluMetSerLysThrAlaLeuGlnPhe 292
RESULT 42	
ST1B1_CANFA	
ID	ST1B1_CANFA STANDARD; PRT; 296 AA.
AC	Q95JD5;
DT	04-JAN-2005, integrated into UniProtKB/Swiss-Prot.
DT	01-DEC-2001, sequence version 1.
DT	07-FEB-2006, entry version 22.
DE	Sulfotransferase family cytosolic 1B member 1 (EC 2.8.2.-)
DE	(Sulfotransferase 1B) (CSULT1B1).
GN	Name=SULT1B1; Synonyms=ST1B2;
OS	Canis familiaris (Dog).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
OC	Canis.
OX	NCBI_TaxID=9615;

[1]	
NUCLEOTIDE SEQUENCE [MRNA], FUNCTION, AND TISSUE SPECIFICITY.	
RP	TISSUE=Liver;
RC	MEDLINE=21261348; PubMed=11368519; DOI=10.1006/abbi.2001.2373;
RX	Tsai C., Falany C.N., Morgenstern R., Swedmark S.;
RA	"Molecular cloning, expression, and characterization of a canine
RT	sulfotransferase that is a human ST1B2 ortholog.";
RL	Arch. Biochem. Biophys. 390:87-92(2001).
CC	-!- FUNCTION: Catalyzes the sulfate conjugation of many hormones,
CC	neurotransmitters, drugs and xenobiotic compounds. Sulfonation
CC	increases the water solubility of most compounds, and therefore
CC	their renal excretion, but it can also result in bioactivation to
CC	form active metabolites. Sulfates thyroid hormones including 3,3'-
CC	diiodothyronine, 3,5,3'-triiodothyronine and 3,5'-diiodothyronine
CC	(lesser degree) and phenols such as 1-naphthol and p-nitrophenol.
CC	-!- SUBUNIT: Monomer (By similarity).
CC	-!- SUBCELLULAR LOCATION: Cytoplasm (By similarity).
CC	-!- TISSUE SPECIFICITY: Expressed highly in the colon, kidney and
CC	small intestine of male and female dogs. Highly expressed in the
CC	jejunum and ileum of the male dog than the female dog, which
CC	displayed more expression in duodenum (at protein level).
CC	-!- SIMILARITY: Belongs to the sulfotransferase family.
CC	Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC	Distributed under the Creative Commons Attribution-NoDerivs License
CC	-----
EMBL	AY004332; AAF86583.1; -; mRNA.
DR	SMR; Q95JD5; 4-296.
DR	Ensembl; ENSCAFG0000002871; Canis familiaris.
DR	InterPro; IPR000863; Sulfotransferase.
DR	Pfam; PF00685; Sulfotransfer_1; 1.
DR	ProDom; PD001218; Sulfotransferase; 1.
KW	Lipid metabolism; Steroid metabolism; Transferase.
FT	CHAIN 1 296 Sulfotransferase family cytosolic 1B
FT	member 1.
FT	/FTID=PRO_0000085160.
FT	NP_BIND 48 53 PAPS (By similarity).
FT	NP_BIND 131 139 PAPS (By similarity).
FT	NP_BIND 194 230 PAPS (By similarity).
FT	ACT_SITE 109 109 Proton acceptor (By similarity).
FT	BINDING 259 259 PAPS (By similarity).
SQ	SEQUENCE 296 AA; 34869 MW; 3D54CF003BF09365 CRC64;
Alignment Scores:	
Pred. No.:	1.63e-22 Length: 296
Score:	443.00 Matches: 101
Percent Similarity:	52.4% Conservative: 51
Best Local Similarity:	34.8% Mismatches: 118
Query Match:	10.1% Indels: 20
DB:	1 Gaps: 5
US-10-768-158-1 (1-2419) x ST1B1_CANFA (1-296)	
QY	51 ACCCGCGGGAGTTCGAGACCAAGTACTTCGAG---TTCCATGCGGTGCGGCTGCGCCC 107
Db	: : : : :
Db	3 SerProLysAspPheLeuArgLysAsnLeuLysMetIleHisGlyTyrProIleIleTyr 22
QY	108 TTCTGCGCGGGAAGATGGAGGAGATCGCAACTTCCCGGTGCGGCCAGCGCTGTGG 167
Db	: : : : :
Db	23 ThrPheAlaAsnAenTrpGluAsnIleGluGlnPheHisSerArgProAspApeIleIle 42
QY	168 ATCGTCACTTACCCCAAGTCCGACCAAGCTTCGTCGAGGAGTGTGTCTACTTGGTGAGC 227
Db	: : : : :
Db	43 IleAlaThrTyrProLysSerGlyThrTrpValSerGluIleValAspMetValLeu 62
QY	228 CAGGCGCTCACCCCGATGATCGGCTTGATGAACATCCAGCAGCAGCTCCCGGTCTG 287
Db	: : : : :
Db	63 AsnAenGlyAspValGluLysCysLysArgAspPheIleThrValLysValProMetLeu 82
QY	288 GAGTACCCACAGCGCGGCGCTG-----GACATCATCAAGGAAGTACCTCT 332
Db	: : : : :
Db	83 GluMetAlaValProGlyLeuArgThrSerGlyIleGlnLeuGluLysAsnProSer 102

```
QY 333 CCCGCCTCATCAAGAGCCACCTGCCCTACCGCTTTCTGCGCTTCTGACCTCCACATGGA 392
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 393 GACTCCAGGTCATCTATATGGCTCGCAACCCCAAGGATCTGGTGTCTTATATCAAG 452
Db ::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 123 AsnCysylsMetIleTyrLeuAlaArgAsnAlaLysAspValAlaValSerTyrTyrHis 142
QY 453 TTCACCGCTCTCGCGGACCATGAGCTACCGAGCCACCTTTCAGAAATTCGCGGAGG 512
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 143 PheAspLeuMetAsnAsnLeuGluProAlaProGlyProTyrGluGluTyrLeuGluArg 162
QY 513 TTTATGAATGATAGCTGGGCTACGGCTCCCTGGTTTGAGACGTCGAGGAGTCTCGGAG 572
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 163 PheMetThrGlyAsnValAlaTyrGlySerTyrPheAsnHisValLysSerTyrTrpLys 182
QY 573 CACCGCATGGACTCGAAGCTGCTTTTCTCAAGTATGAAGACATCGCATCGGACCTGGTG 632
Db ::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 183 LysLysGluGluHisProIleLeuPheLeuTyrTyrGluAspMetLysGluAsnProLys 202
QY 633 ACGATGTGTGAGACGCTGGCCAGATTCCTGGGGGTGCTCTGTGACAAGGCCAGCTGGAA 692
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 203 ArgGluValGlnLysIleAlaArgPheLeuGluLysAsnLeuAsnAspGluValLeuAsp 222
QY 693 GCCCTGACGGAGCACTGC-----CACGAGCTGGTGACACAG---TGCTGCAACGCTGAG 743
Db ::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 223 LysIleIleHisHisThrSerPheGluMetMetLysAspAsnProLeuValAsnTyrThr 242
QY 744 GCCTTGCC-----GTGGCCCGGGGAAGAGTT 770
Db ::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 243 HisLeuProSerThrValMetAspHisSerLysSerPheMetArgLysGlyIleAla 262
QY 771 GCGCTGTGGAAGGACATCTTCACGCTCTCCATGATGAGAAGTTTGACTGGTGATATAA 830
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 263 GlyAspTrpLysAsnTyrPheThrValAlaGlnAsnGluLysPheAspValIleTyrLys 282
QY 831 CAGAAGATGGAAAGTGTGACCTCACGTTT 860
Db ::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 283 LysGluMetSerGlyThrThrLeuGlnPhe 292

RESULT 43
STIC1_RABIT STANDARD; PRT; 296 AA.
AC O46503;
DT 01-DEC-2000, integrated into UniProtKB/Swiss-Prot.
DT 01-JUN-1998, sequence version 1.
DT 07-FEB-2006, entry version 33.
DE Sulfotransferase 1C1 (EC 2.8.2.-) (rabSULT1C2).
GN Name=SULT1C1;
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Lagomorpha; Leporidae;
OC Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP NUCLEOTIDE SEQUENCE [mRNA], AND CHARACTERIZATION.
RC TISSUE=Stomach;
RA MEDLINE=99410866; PubMed=10481272; DOI=10.1016/S1357-2725(99)00038-2;
RA Hehonian N., Zhu X., Brix L., Bolton-Grob R., Barnett A., Windmill K.,
RA McManus M.;
RT "Molecular cloning, expression, localisation and functional
RT characterisation of a rabbit SULT1C2 sulfotransferase.";
RL Int. J. Biochem. Cell Biol. 31:869-882(1999).
CC -!- FUNCTION: Catalyzes the sulfate conjugation of many drugs,
CC xenobiotic compounds, hormones, and neurotransmitters. Shows
CC activity towards p-nitrophenol. Probably the major stomach
CC sulfotransferase.
CC -!- SUBCELLULAR LOCATION: Cytoplasm.
CC -!- TISSUE SPECIFICITY: Found in gastrointestinal tract tissues, liver
CC and kidney.
CC -!- SIMILARITY: Belongs to the sulfotransferase family.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
```

```
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF026304; AAC00410.1; -, mRNA.
DR HSSP; P50224; ICJM.
DR SMR; O46503; 12-296.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransfer 1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Transferase.
FT CHAIN 1 296 Sulfotransferase 1C1.
FT NP_BIND 49 54 /FTID=PRO_0000085134.
FT NP_BIND 131 139 PAPS (By similarity).
FT NP_BIND 194 230 PAPS (By similarity).
FT NP_BIND 258 260 PAPS (By similarity).
FT ACT_SITE 109 109 Proton acceptor (By similarity).
SQ SEQUENCE 296 AA; 34572 MW; 4A1089D0F8D4F295 CRC64;

Alignment Scores:
Pred. No.: 1-77e-22 Length: 296
Score: 442.50 Matches: 95
Percent Similarity: 56.1% Conservative: 66
Best Local Similarity: 33.1% Mismatches: 107
Query Match: 10.1% Indels: 19
DB: 1 Gaps: 6

US-10-768-158-1 (1-2419) x STIC1_RABIT (1-296)
QY 54 CCGGGGAGTTCCAGAGCAAGTACTTCGATTCATTCGCGTGGCGTCCGCTTCTGTC 113
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 114 CCGGGGAGATGAGGAGATCGCAACTTCCCGGTGGCGCCAGCACGTCGTGGATGTC 173
Db ::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 174 ACCTACCCCAAGTCCGCGACACGCTTGTCTCAGAGGTGCTTACTTGTGTGAGCCAGGC 233
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 234 GCTGACCCCGATGAGATCGGCTTGTGTAACATCGACGAGCAGCTCCCGTCTCGGAGTAC 293
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 294 -----CCAGCGCG---GGCTGGACATCATCAAGAACTGACCTCCCGCCTC 341
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 342 ATCAAGACCCACCTCGCCTACCGCTTCTGCGCTTCTGACCTCCCAATGAGACTCCAAG 401
Db ::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 402 GTCATCTATATGCTCGCAACCCCAAGGATCTCGTGTGTCTTATATCATGTTCCACCGC 461
Db ::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 462 TCTCTCGGACCATGAGTACCGAGCAGCTTTCAGAAATTCGCGGAGGTTTATGAAT 521
Db ::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 522 GATAAGCTGGGCTACGGCTCTCGTGTGACAGCTGACGAGTCTGGGAGCAGCCGATG 581
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 582 GATCGAAGCTGCTTTTCTCAAGTATGAAGATGATGATCGGACCTGCTGTGAGATGGTG 641
Db ::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 642 GAGCAGCTGCCAGATTCCTCGGGGTGCTCTGTGACAAGGCCCGGCTGGAAGCCCTGACG 701
Db ::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 702 GAGCACTGC-----CACGAGCTGGTGGAGCAG---TGCTGCAACGCTGAGGCCCTGCC 752
```

```

Db 226 GlnGluThrSerPheGluLysMetLysAspAsnProMetIleAsnArgSerThrValPro 245
Qy 753 -----GTGGCCGGGGAAGAGTTGGGCTGGG 779
Db 246 LysSerIleMetAspGlnSerIleSerProPheMetArgLysGlyThrValGlyAspTrp 265
Qy 780 AAGGACATCTTCACGCTCCCATGAATGAGAGTTGTGATGGTGATATAACAGAGATG 839
Db 266 LysAsnHisPheThrValAlaGlnSerHisArgLeuAspGluLeuTyrArgLysLysMet 285
Qy 840 GGAAAGTGTGACCTCACGTTT 860
Db 286 GluGlyValSerIleAspPhe 292

RESULT 44
ID Q9XT99_RABIT PRELIMINARY; PRT; 292 AA.
AC Q9XT99;
DT 01-NOV-1999, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, entry version 17.
DE Arylsulfotransferase S1A8.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Lagomorpha; Leporidae;
OC Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=New Zealand White; TISSUE=Liver;
RX MEDLINE=2109889; PubMed=11181495;
RA Honma W., Kamiyama Y., Yoshinari K., Sasano H., Shimada M., Nagata K.,
RA Yamazoe Y.;
RT "Enzymatic characterization and interspecies difference of phenol
RT sulfotransferases, S1A forms.";
RL Drug Metab. Dispos. 29:274-281 (2001).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AB029494; BAA82295.1; -, mRNA.
DR HSSP; P50224; 1CUM.
DR SMR; Q9XT99; 8-289.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransfer_1;
DR ProDom; PD001218; Sulfotransferase; 1.
KW Transferase.
SQ SEQUENCE 292 AA; 33841 MW; 2BF5F63A259DFFD1 CRC64;

Alignment Scores:
Pred. No.: 1.92e-22 Length: 292
Score: 442.00 Matches: 95
Percent Similarity: 53.9% Conservative: 50
Best Local Similarity: 35.3% Mismatches: 104
Query Match: 10.0% Indels: 20
DB: 2 Gaps: 4

US-10-768-158-1 (1-2419) x Q9XT99_RABIT (1-292)
Qy 90 GGCGTGGCGTCCCGCTCTCGCGGGGAAGATGGAGGATGCGCAACTTCCCGGTG 149
Db 17 GlyValProLeuIleLysTyrPheAlaGluAlaMetGlyProLeuGlnSerPheLysAla 36
Qy 150 CGGCCACCGACGTTGGATCGTACCTACCCAGTCCCGACCGACGCTCTCTCGAGGAG 209
Db 37 ArgProAspAspLeuLeuIleSerThrTyrProLysSerGlyThrThrTipValSerGlu 56
Qy 210 GTGGTCTACTTGGTGGACCGGCGCTGACCCCGATGATGATGCGGCTTGATGAATCGAC 269
Db 57 IleLeuAspMetIleTyrGlnSerGlyAspGlnGlnLysCysLeuArgAlaProIleTyr 76

```

```

Qy 270 GAGCAGCTCCCGTCTCTGGAGTACCCACAGCCG-----GCCCTGGACATCATC 317
Db 77 IleArgValProPheLeuGluPheLysAlaProSerGlyMetGluThrLeu 96
Qy 318 AAGGAATGACCTCTCCCGCTCATCAAGAGCCACCTGCGCTACCGCTTCTGCCCTCT 377
Db 97 LysAspThrProSerProArgLeuLeuLysThrHisLeuProLeuAlaLeuLeuProGln 116
Qy 378 GACCTCCACAAATGAGACTCCAAGTCTATATGCTCGCAACCCCAAGATCTGGTG 437
Db 117 ThrLeuLeuAspGlnLysValLysValValTyrValAlaArgAsnAlaLysAspValAla 136
Qy 438 GTGTCTTATTATCAGTTCCACCGCTCTCTCGGACCATAGCTACCGAGGACCTTCAA 497
Db 137 ValSerTyrTyrAsnPheTyrHisMetAlaArgValHisProGluProGlyThrTrpAsp 156
Qy 498 GAATTCCTCCGGAGGTTTATGAATGATAAGCTGGCTACGCTCGCTCTGTTTGACGACGTG 557
Db 157 SerPheLeuGluLysPheMetAlaGlyLysValSerTyrGlySerTyrTyrGlnHisVal 176
Qy 558 CAGGAGTTCTGGGAGACCCGCTGCTCGAAGCTGCTTTTCTCAAGTATGAAGACATG 617
Db 177 ArgGluTrpTrpGluLeuSerArgThrHisProValLeuTyrLeuPheGluAspMet 196
Qy 618 CATCGGACCTGTGTGACGATGGTGAGCAGCTGGCCAGATTCCTGGGGGTGCTCTGTGAC 677
Db 197 LysGluAsnProLysThrGlnIleLysLysIleLeuGluPheLeuGlyArgSerLeuPro 216
Qy 678 AAGGCCAGCTGGNAGCCCTGACGAGCAGCTGCCAGCTGGTGGGACCAG----- 728
Db 217 GluGluThrValAspArgIleAla---HisCysThrSerPheLysGluMetLysLysAsn 235
Qy 729 ---TGCTGCAACGCTGAGGCCCTGCC----- 752
Db 236 ProMetThrAsnTyrSerThrIleProGluAsnIleMetAspHisAsnValSerProPhe 255
Qy 753 GTGGCGCGGGAAGAGTTGGCTGTGGAAGGACATCTTACCGCTCCATGAATGAGAG 812
Db 256 MetArgArgGlyValAlaGlyAspTyrLysThrThrPheThrValAlaGlnAsnGluArg 275
Qy 813 TTTGACTTGGTGATATAAACAGAGATG 839
Db 276 PheGluAlaAspTyrAlaGluLysMet 284

RESULT 45
Q6GP49_XENLA PRELIMINARY; PRT; 297 AA.
ID Q6GP49_XENLA PRELIMINARY; PRT; 297 AA.
AC Q6GP49;
DT 19-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, entry version 14.
DE MGC80677 protein.
GN Name=MGC80677;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Spleen;
RX MEDLINE=23388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hong L.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

```

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences".
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Spleen;
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.;
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 RT initiative".
 RL Dev. Dyn. 225:384-391 (2002).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Spleen;
 RA Klein S., Gerhard D.S.;
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC -----
 DR EMBL; BC073295; AAH73295.1; -; mRNA.
 DR GO; GO:0008146; F:sulfotransferase activity; IEA.
 DR InterPro; IPR000863; Sulfotransferase.
 DR Pfam; PF00685; Sulfotransfer_1; 1.
 DR ProDom; PD001218; Sulfotransferase; 1.
 DR PROSITE; PS00001; Sulfotransferase; 1.
 SQ SEQUENCE 297 AA; 34860 MW; 88D37337F60C6603 CRC64;

Alignment Scores:

Pred. No.: 2,46e-22 Length: 297
 Score: 440.50 Matches: 99
 Percent Similarity: 52.1% Conservativity: 51
 Best Local Similarity: 34.4% Mismatches: 112
 Query Match: 10.0% Indels: 26
 DB: 2 Gaps: 6

US-10-768-158-1 (1-2419) x Q6GP49_XENLA (1-297)

Qy 65 CGAGAGCAAGTACTTCGAGTTCCATGGCGTGGCGCTGCGGCCTTCTGCGC----- 116
 Db 10 ArgPheGlnLeuLeuProVal-HisGlyIle-----ProPheMetLysProIleAl 26
 Qy 117 -GGGAAGATGAGGAGATCGCAACTTCCCGTGGCGCCAGCGAGTGTGGTGCATCGTCAC 175
 Db 26 aAspAsnTrpAspArgIleGluThrPheGlnAlaLysProAspAspLeuIleAlaTh 46
 Qy 176 CTACCCCAAGTCCGGCACCAGCTTGTCTGAGGAGTGTCTACTTGTGTGAGCCAGGCGC 235
 Db 46 rTyrProLysAlaGlyThrTrpMetGlnGluIleValAspSerIleMetAsnAlaG1 66
 Qy 236 TGACCCCGATGAGTGGCTTGATGAACATCGACGAGCAGCTCCCGTCTCTGGAGTAC-- 293
 Db 66 yAspLeuLysLysValLysArgAlaProThrHisValArgPheProPheLeuGluIleCy 86
 Qy 294 ----CCACAGCG-----GCCTGGACATCATCAAGGAACGTACCTCTCCCGCCT 340
 Db 86 sAsnProSerProValProCysGlyValAspValIleGluIleThrGlnSerProArgil 106
 Qy 341 CATCAAGAGCCACCTGCCCTACCGCTTCTGCGCCTCTGACCTCCACAAATGAGACTCCAA 400
 Db 106 eileLysThrHisLeuProTyrGluLeuIleProLysSerPheTrpGluHisGlyCysLy 126
 Qy 401 GGTATCTATATGCTCGCACCCCAAGGATCTGGTGTCTTATATCATAGTTCCACCG 460
 Db 126 sValIleTyrValAlaArgAsnAlaLysAspAsnAlaValSerTyrTyrPheAspLe 146

Qy 461 CTCTCTCGGACCATGAGCTACCGAGGCACCTTTCAAGAAATCTGCGGAGGTTATGAA 520
 Db 146 uMetAsnLysThrGlnProHisProGlyThrTrpGluGluTyrValGlyLysPheLeuLy 166
 Qy 521 TGATAAGCTGGGTACGGCTCCCTGGTTTGAGCAGCTGACGAGTTCCTGGAGCACCACAT 580
 Db 166 sGlyAsnValProTrpGlyTrpPheHisHisValIleGlyTrpTrpLysAlaArgAl 186
 Qy 581 GGACTCGAAGCTGCTTTTCTCAAGTATGAAGACATGCATCGGAGCCTGGTGCAGGTGT 640
 Db 186 aLysHisGlnIleLeuTyrValPheTyrGluAspMetLysGluAspProLysArgGluI1 206
 Qy 641 GGAGCAGCTGCCAGATTCCTGGGGTGTCTGTGACAAAGCCCGCAGCTGGAAGCCCTGAC 700
 Db 206 eArgLysValIleArgPheLeuGlyLysGluLeuSerGluAspLeuGluLysIleCy 226
 Qy 701 GGAGCAGCTGCCAC-----CAGCTGGTGGACAGTGTCTGCAAGCTGAGGCGCTGCC 751
 Db 226 sGlnHisThrSerPheLysAlaMetLysGluAsnProMetAlaAsnTyrSerAlaMetPr 246
 Qy 752 C-----GTGGCGCGGGAAGAGTTGGCGTGTG 778
 Db 246 oAspSerValLeuAspGlnSerIleSerProPheMetArgLysGlyGluValSerAspTr 266
 Qy 779 GAAGGACATCTTCACCGTCTCCATGAATGAGAGATTGACTTGGTGTATAAACAGAAAGAT 838
 Db 266 pLysAsnHisPheThrValGlnGlnAsnGluMetPheAspAlaGluTyrGlnLysArgMe 286
 Qy 839 GGGAAAGTGTGACCTCACGTTT 860
 Db 286 tGluGlyThrAspLeuLysPhe 293
 RESULT 46
 Q3TOY3_BOVIN
 ID Q3TOY3_BOVIN PRELIMINARY; PR; 296 AA.
 AC Q3TOY3;
 DT 11-OCT-2005, integrated into UniProtKB/TrEMBL.
 DT 11-OCT-2005, sequence version 1.
 DT 07-MAR-2006, entry version 4.
 DE Hypothetical protein.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
 OC Pecora; Bovidae; Bovinae; Bos.
 OC NCBI_TaxID=9913;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=Crossbred x Angus; TISSUE=ileum;
 RA Moore S., Alexander L., Brownstein M., Guan L., Lobo S., Meng Y.,
 RA Tanaguchi M., Wang Z., Yu J., Prange C., Schreiber K., Shenmen C.,
 RA Wagner L., Bala M., Barbazuk S., Barber S., Babakiaff R., Beland J.,
 RA Chun E., Del Rio L., Gibson S., Hanson R., Kirkpatrick R., Liu J.,
 RA Matsuo C., Mayo M., Santos R.R., Stott J., Teal M., Wong D.,
 RA Siddiqui A., Holt R., Jones S.J., Marra M.A.;
 RL Submitted (AUG-2005) to the EMBL/GenBank/DBJ databases.
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC -----
 DR EMBL; BC102208; AA102209.1; -; mRNA.
 DR GO; GO:0008146; F:sulfotransferase activity; IEA.
 DR InterPro; IPR000863; Sulfotransferase.
 DR Pfam; PF00685; Sulfotransfer_1; 1.
 DR ProDom; PD001218; Sulfotransferase; 1.
 DR Hypothetical protein.
 SQ SEQUENCE 296 AA; 34674 MW; 8A44653780078275 CRC64;
 Alignment Scores:
 Pred. No.: 4,72e-22 Length: 296
 Score: 436.50 Matches: 94
 Percent Similarity: 53.1% Conservativity: 53
 Best Local Similarity: 33.9% Mismatches: 111
 Query Match: 9.9% Indels: 19

```
DB: 2 Gaps: 4
US-10-768-158-1 (1-2419) x Q3T0Y3_BOVIN (1-296)
87 CATGCGCTGGCGCTCCGCGCTTCCTGCGCGGGAAGATGGAGGAGATCGCAACTTCCCG 146
16 HisGlyCysProIleThrTyralaPhealaAsnAsnTrpGluLysIleGluGlnPheGln 35
147 GTGCGGCGCCAGCGAGCTGTGATCGTACCTACCCCAAGTCCGGCAGCAGCTTGTGCGAG 206
36 SerArgProAspAspIleMetIleValThrTyProLysSerGlyThrThrTriPleSer 55
207 GAGGTGGTCTACTGGTGAGCAGCGCGCTGACCCCGATGAGATCGGCTTGATGACATC 266
56 GluIleValAspMetValLeuHisAspGlyAspValGluLysCysLeuArgAspValIle 75
267 GACGAGCAGCTCCCGGTCTGAGTACCCACACGCG-----GGCTGGAC 311
76 ThrAlaLysValProMetLeuGluLeuAlaLeuProGlyLeuArgThrSerGlyLeuGlu 95
312 ATCATCAAGGAATGACCTCTCCCGCTCATCAAGAGCCACTGCTCCCTACCGCTTTCTG 371
96 GlnLeuGluLysAsnProSerProArgValValLysThrHisLeuProIleAspLeuIle 115
372 CCCTCTGACCTCCACAATGGAGACTCCCAAGTCACTATATGCTGCGCAACCCCAAGAT 431
116 ProLysSerPheTrpGluAsnAsnCysLysIleIleTyLeuAlaArgAsnAlaLysAsp 135
432 CTGTTGGTGTCTATTATCAGTTCCACCGCTCTCTGCGGACCATGAGCTACCGAGGCACC 491
136 ValAlaValSerPheTyHisPheAspLeuMetAsnLeuGlnProLeuProGlyThr 155
492 TTCAAGAATTTCTGCGGAGGTTTATGAATGATAAGCTAGGCTCGGCTCTCTGTTTGCAG 551
156 TrpGlyGluTyLeuGluLysPheLeuThrGlyAsnValAlaTyGlySerTrpPheAsn 175
552 CAGCTGCGAGGAGTCTGGGAGCAGCGGATGAGTCAAGCTGCTTTTCTCAAGTATGAA 611
176 HisValLysSerTrpTrpLysLysLysGluGlyHisProIleLeuPheLeuPheTyGlu 195
612 GACATGATCGGACCTGTCAGCATGTCGAGCAGCTGGCAGATTCCTGGGGTGTC 671
196 AspMetLysGluAsnProLysGlnGluIleLysLysValValArgPheLeuGluLysAsn 215
672 TGTGACAAAGGCCAGCTGGAAGCCCTGAGGAGCACTGC-----CACCAGCTGTGTCAG 725
216 LeuAspAspGluIleLeuAspLysIleIleTyHisSerPheGluMetMetLysAsp 235
726 CAG---TGCTGCAAGCTGAGCGCTGCGC----- 752
236 AsnProLeuValAsnTyThrHisLeuProSerGluValMetAspHisSerLysSerSer 255
753 ---GTGGCGCGGAGAGTGGGCTGTGGAGGACATCTTCACGCTCCATCAATGACAG 809
256 PheMetArgLysGlyIleAlaGlyAspTrpLysAsnTyPheThrValAlaGlnAsnGlu 275
810 AAGTTTGACTGGTGATATAAACAAGATGGGAAGTGTGCCTCACCTT 860
276 LysPheAspAlaIleTyLysLysGluMetSerGluThrGluLeuGlnPhe 292
RESULT 47
ST1E2 RAT
ID ST1E2 RAT STANDARD; PRT; 295 AA.
AC P52845;
DT 01-OCT-1996, integrated into UniProtKB/Swiss-Prot.
DT 01-OCT-1996, sequence version 1.
DE 07-FEB-2006, entry version 38.
DE Estrogen sulfotransferase, isoform 2 (EC 2.8.2.4) (EST-2)
DE (Sulfotransferase, estrogen-prefering) (Estrogen sulfotransferase).
GN Name=Ste2; Synonyms=Ste;
OS Rattus norvegicus (Rat);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
```

```
OC Muroidea; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN NUCLEOTIDE SEQUENCE [MRNA].
RP STRAIN=Fischer 344; TISSUE=Liver;
RX MEDLINE=96305357; PubMed=8688469; DOI=10.1016/0167-4781(96)00065-6;
RA Rikke B.A., Roy A.K.;
RT "Structural relationships among members of the mammalian
RL sulfoftransferase gene family.";
RL Biochim. Biophys. Acta 1307:331-338(1996).
CC -!- FUNCTION: Sulfation of estrone and estradiol. May control the
CC level of the estrogen receptor by sulfurylating free estradiol (By
CC similarity).
CC -!- CATALYTIC ACTIVITY: 3'-phosphoadenylyl sulfate + estrone =
CC adenosine 3',5'-bisphosphate + estrone 3-sulfate.
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasm (By similarity).
CC -!- SIMILARITY: Belongs to the sulfotransferase family.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC EMBL: U50205; AAB07681.1; -; mRNA.
DR HSSP; P49891; 1306.
DR SMR; P52845; 7-294.
DR Ensembl; ENSRNOG00000001957; Rattus norvegicus.
DR RGD; 3776; Ste.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase 1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Lipid-binding; Steroid-binding; transferase.
FT CHAIN 1 295 Estrogen sulfotransferase, isoform 2.
FT NP_BIND 48 53 PAPS (By similarity).
FT NP_BIND 130 138 PAPS (By similarity).
FT NP_BIND 193 229 PAPS (By similarity).
FT NP_BIND 257 259 PAPS (By similarity).
FT ACT_SITE 108 108 Proton acceptor (By similarity).
SQ SEQUENCE 295 AA; 35365 MW; 149B5C9D46039AAF CRC64;
Alignment Scores:
Pred. No.: 6,03e-22 Length: 295
Score: 435.00 Matches: 92
Percent Similarity: 53.3% Conservative: 60
Best Local Similarity: 32.3% Mismatches: 115
Query Match: 9.9% Indels: 18
DB: 1 Gaps: 3
US-10-768-158-1 (1-2419) x ST1E2_RAT (1-295)
Qy 60 GAGTTCGAGAGCAAGTACTTCGAGTTCATCGCGCTGCGCTGCGCGCTTCTGCGCGGG 119
Db 7 GluTyTrpGluValPheGlyAspPheHisGlyPheLeuMetAspLysLeuPheThrLys 26
Qy 120 AAGATGAGGAGATCGCAACTTCCCGGTGCGCGCGCGCGAGCGTGTGATCGTCACTAC 179
Db 27 TyrTrpGluAspValGluThrPheSerAlaArgProAspAspLeuLeuValThrTyr 46
Qy 180 CCCAAGTCCCGCACCACTGCTGTCAGAGTGCTTACTTGTGGTGAGCGCGCGCTGAC 239
Db 47 ProLysSerGlySerThrTrpIleGlyGluIleValAspMetIleTyLysGluGlyAsp 66
Qy 240 CCCGATGAGATCGGCTTGATGAACATCGACGAGCAGCTCCCGGTCTCCGGAGTACCCACAG 299
Db 67 ValGluLysCysLysGluAspAlaIlePheAsnArgIleProTyLeuGluCysArgAsn 86
Qy 300 CCG-----GGCTGGAGCATCATCAAGAACTGACCTTCCCGCTCATCAAG 347
Db 87 GluAspLeuIleAsnGlyIleLysGlnLeuLysGluLysGluSerProArgIleValLys 106
Qy 348 AGCCACCTGCGCTACCGCTTTCGCTCTGACCTCCACCAATGGAGACTCCAGGTGATC 407
Db 107 ThrHisLeuProAlaLysLeuLeuProAlaSerPheTrpGluLysAsnCysLysIleIle 126
```

QY 408 TATATGGCTCGCACCCAGGATCTGGTGTGCTTATTATTCAGTTCCACCGCTCTCTG 467
 Db 127 TyrLeuCysArgAsnAlaLeuAspValValSerTyrTyrPhePheLeuLeuMet 146
 QY 468 CGGACCATGAGCTACCGAGGACCTTCAAGAAATCTGCGGAGGTTTATGAATGATAAG 527
 Db 147 LysSerTyrProAnProLysSerPheSerGluPheValGluLysPheMetGluGluGln 166
 QY 528 CTGGGCTACGGCTCTCTGGTTAGCACGTGCAGAGTCTTGGGAGCACCGCATGACTCG 587
 Db 167 ValProTyrGlySerTyrTyrAspHisValLysSerTrpTrpGluLysSerLysAsnSer 186
 QY 588 AACCTGCTTTTCTCAAGTATGACATGATCGGACCTGCTGAGTGGTGAGGACGAG 647
 Db 187 ArgValLeuPheMetPheTyrGluAspMetLysGluAspMetLysGluValValLys 206
 QY 648 CTGGCCAGATTCTCGGGGTCTCTGTGACAAAGCCAGCTGGAAGCCCTGACGAGCAC 707
 Db 207 LeuLeuGluPheLeuGluArgAspProSerAlaGluLeuValAspArgIleIleGlnHis 226
 QY 708 TGC-----CACACAGCTGGTGGACAGTGTGCAACGCTGAGGCCCTGCC----- 752
 Db 227 ThrSerPheGlnGluMetLysAsnAsnProCysThrAsnTyrSerMetLeuProGluThr 246
 QY 753 -----GTGGCCGGGGAAGAGTTGGCTGTGGAAGGAC 785
 Db 247 MetLeuAspLeuLysValSerProPheMetArgLysGlyLeuValGlyAspTrpLysAsn 266
 QY 786 ATCTTCCCGCTCTCCATGAATGAGAAGTTTGACTGTGTATTAACAGAGATGGGAAG 845
 Db 267 HisPheProGluAlaLeuArgGluArgPheGluGluHisTyrGlnGlnMetLysAsp 286
 QY 846 TGTGACCTCAGTTT 860
 Db 287 CysProValLysPhe 291

RESULT 48
 ST16 RAT
 ID ST16 RAT STANDARD; PRT; 295 AA.
 AC P49890;
 DT 01-OCT-1996, integrated into UniProtKB/Swiss-Prot.
 DT 01-OCT-1996, sequence version 1.
 DT 07-FEB-2006, entry version 37.
 DE Estrogen sulfotransferase, isoform 6 (EC 2.8.2.4) (EST-6)
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridea; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [MRNA].
 RC TISSUE=Liver;
 RX MEDLINE=95161323; PubMed=7857871; DOI=10.1016/0960-0760(94)00147-5;
 RA Falany J.L., Krasnykh V., Mikheeva G., Falany C.N.;
 RT "Isolation and expression of an isoform of rat estrogen
 sulfotransferase.";
 RL J. Steroid Biochem. Mol. Biol. 52:35-44 (1995).
 CC -1- FUNCTION: Sulfation of estrone and estradiol. May control the
 level of the estrogen receptor by sulfurylating free estradiol.
 CC -1- CATALYTIC ACTIVITY: 3'-phosphoadenylyl sulfate + estrone =
 CC adenosine 3',5'-bisphosphate + estrone 3-sulfate.
 CC -1- SUBUNIT: Homodimer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasm.
 CC -1- DEVELOPMENTAL STAGE: Expressed in males rats, but not in females.
 CC -1- INDUCTION: Induced by androgens and suppressed by estrogens.
 CC Expression is under the influence of pituitary growth hormone and
 thyroid hormone.
 CC -1- SIMILARITY: Belongs to the sulfotransferase family.
 CC
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License

CC -----EMBL; S76490; AAB33442.1; -; mRNA.
 DR HSP; P49891; IBO6.
 DR SMR; P49890; 7-294.
 DR Ensembl; ENSRNOG00000001957; Rattus norvegicus.
 DR InterPro; IPR000863; Sulfotransferase.
 DR Pfam; PF00685; Sulfotransferase 1; 1.
 DR ProDom; PD001218; Sulfotransferase; 1.
 KW Lipid-binding; Steroid-binding; Transferase.
 FT CHAIN 1 295 Estrogen sulfotransferase, isoform 6.
 FT PAPS (By similarity).
 FT NP_BIND 48 53 PAPS (By similarity).
 FT NP_BIND 130 138 PAPS (By similarity).
 FT NP_BIND 193 229 PAPS (By similarity).
 FT NP_BIND 257 259 PAPS (By similarity).
 FT ACT_SITE 108 108 Proton acceptor (By similarity).
 SQ SEQUENCE 295 AA; 35303 MW; 9FBD5861AFDC9B05 CRC64;
 Alignment Scores:
 Pred. No.: 6.03e-22 Length: 295
 Score: 435.00 Matches: 93
 Percent Similarity: 53.1% Conservative: 62
 Best Local Similarity: 31.8% Mismatches: 105
 Query Match: 9.9% Indels: 32
 DB: 1 Gaps: 4
 US-10-768-158-1 (1-2419) x ST16 RAT (1-295)
 QY 60 GAGTTCGAGACCAAGTACTTCGAGTTCCATGGCGTGGCTGCCGCCCTTCTCGCGGG 119
 Db 7 GluTyrTyrGluValPheGlyAspPheHisGlyValLeuValAspLysLeuPheThrLys 26
 QY 120 AAGATGAGAGAGATCGCAACTCCCGGTGCGGCCAGCAGACGTGGATCGTCACTAC 179
 Db 27 TyrTrpGluAspValGluThrPheSerAlaArgProAspAspLeuValThrTyr 46
 QY 180 CCAAGTCCGGCACCGACTTGTGTCAGGAGTGTCTACTTGTGAGCCAGCGCGTGAC 239
 Db 47 ProLysSerGlySerThrTrpIleGlyLeuValAspMetIleTyrLysGluGlyAsp 66
 QY 240 CCGATGAGATCGGCTTGATGAACATCGACGACGATCCCGTCTCTGGAGTACCCACAG 299
 Db 67 ValGluLysCysLysGluAspAlaIlePheAsnArgIleProTyrLeuGluCysArgAsn 86
 QY 300 CCG-----GGCTGGACATCATCAAGAACTGACCTCTCCCGCCTCATCAAG 347
 Db 87 GluAspLeuLeuAsnGlyIleLysGlnLeuLysGluSerProArgIleValLys 106
 QY 348 AGCACCTGCGCTACCGCTTCTGCGCTCTGACCTCCACAATGGAGACTCCAAAGTCTATC 407
 Db 107 ThrHisLeuProAlaLysLeuLeuProAlaSerPheTrpGluLysAsnCysLysIle 126
 QY 408 TATATGGCTCGCAACCCCAAGGATCTGGTGTGCTTATTATTCAGTTCCACCGCTCTCTG 467
 Db 127 TyrLeuCysArgAsnAlaLysAspValValSerTyrTyrPhePheLeuLeuIle 146
 QY 468 CGGACCATGAGCTACCGAGCACCCTTCAAGAAATCTGCGGAGGTTTATGAATGATAAG 527
 Db 147 LysSerTyrProAnProLysSerPheSerGluPheValGluLysPheMetGluGluGln 166
 QY 528 CTGGGCTACGGCTCTCTGGTTAGCACGTGCAGAGTCTTGGGAGCACCGCATGACTCG 587
 Db 167 ValProTyrGlySerTyrTyrAspHisValLysSerTrpTrpGluLysSerLysAsnSer 186
 QY 588 AACCTGCTTTTCTCAAGTATGACATGATCGGACCTGCTGAGTGGTGAGGACGAG 647
 Db 187 ArgValLeuPheMetPheTyrGluAspMetLysGluAspMetLysGluValValLys 206
 QY 648 CTGGCCAGATTCTCGGGGTCTCTGTGACAAAGCCAGCTGGAAGCCCTGACGAGCAC 707
 Db 207 LeuLeuGluPhe-----GluArgAspProLeuAlaGluLeu 219
 QY 708 TGCCACCGAGCTGTGTG-----GACCAGTGTGCAAC 737

Ebert L., Schick M., Neubert P., Schatten R., Henze S., Korn B.;
 "Cloning of human full open reading frames in Gateway(TM) system entry
 vector [pDONR201].";
 RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
 CC
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC -----
 DR EMBL; CR407621; CAG28549.1; -; mRNA.
 DR SMR; Q53X91; 3-292.
 DR Ensembl; ENSG0000109193; Homo sapiens.
 DR GO; GO:0008146; F:sulfotransferase activity; IEA.
 DR InterPro; IPR000863; Sulfotransferase.
 DR Pfam; PF00685; Sulfotransferase_1; 1.
 DR ProDom; PD001218; Sulfotransferase; 1.
 DR NON_TER 294 294
 SQ SEQUENCE 294 AA; 35126 MW; 9EC8923D20757D57 CRC64;

Alignment Scores:

Pred. No.:	7,1e-22	Length:	294
Score:	424.00	Matches:	95
Percent Similarity:	53.0%	Conservative:	56
Best Local Similarity:	33.3%	Mismatches:	116
Query Match:	9.9%	Indels:	18
DB:	2	Gaps:	3

US-10-768-158-1 (1-2419) x Q53X91_HUMAN (1-294)

Qy	60	GAGTTCGAGGAGCACTTCGAGTTCATGGCTGGCGGTGGCGCCCTCTCGCGGGG	119
Db	6	AspTyrTyrGluLysPheGluValHisGlyIleLeuMetTyrLysAspPheValLys	25
Qy	120	AAGATCGAGGAGATGCCAACTTCGCGTGGCGCCAGCGAGTGTGATCGTCACTAC	179
Db	26	TyrTrpAsnValGluAlaPheGlnAlaArgProAspPheValLleAlaThrTyr	45
Qy	180	CCCAAGTCCGGCACCAGCTTGTGCGAGGAGTGTCTACTTTGGTAGCCAGGCGGTGAC	239
Db	46	ProLysSerGlyThrThrTrpValSerGluIleValTyrMetIleTyrLysGluGlyAsp	65
Qy	240	CCCGATGAGATCGGCTTGATGAACATCGACGAGCTCCGGTCTGGAGTAGTCCACAG	299
Db	66	ValGluLysCysLysGluAspValIlePheAsnArgIleProPheLeuGluCysArgLys	85
Qy	300	CCG-----GGCTGGACATCATCAAGAACTGACCTCTCCCGCCCTCATCAAG	347
Db	86	GluAsnLeuMetAsnGlyValLysGlnLeuAspGluMetAsnSerProArgIleValLys	105
Qy	348	AGCACCTGCGCTACCGCTTCTCGCCCTCTGACCTCCCAATGGAGACTCGAAGTTCATC	407
Db	106	ThrHisLeuProGluLeuLeuProAlaSerPheTrpGluLysAspCysLysIleIle	125
Qy	408	TATATGGCTCGCAACCCCAAGGATCTGGTGGTGTCTTATTATCATGTTCCACCGCTCTCTG	467
Db	126	TyrLeuCysArgAsnAlaLysAspValAlaValSerPheTyrTyrPhePheLeuMetVal	145
Qy	468	CGGACCATGAGCTACCGAGGCACCTTCAAGAAATTCGCGAGGTTTATGAATGATAAG	527
Db	146	AlaGlyHisProAsnProGlySerPheProGluPheValGluLysPheMetGlnGln	165
Qy	528	CTGGGTACGGTCTCGGTTTCAGACCTGACGAGGAGTTCTCGGAGCACCAGCTGCTCG	587
Db	166	ValProTyrGlySerTrpTyrLysHisValLysSerTrpTrpGluLysGlyLysSerPro	195
Qy	588	AACGTGCTTTTCTCAAGTATGAAGCATGCATCGGACCTGGTGGACGATGGTGGAGCAG	647
Db	186	ArgValLeuPheLeuPheTyrGluAspLeuLysGluAspIleArgLysGluValIleLys	205
Qy	648	CTGCCAGATTCTGGGGGTGCTCTGTGCACAGGCCAGCTGGAAGCCCTGACGGAGCAC	707
Db	206	LeuIleHisPheLeuGluArgLysProSerGluGluLeuValAspArgIleHisHis	728
Qy	708	TGCCAC-----CAGCTGGTGGACGAG	728

Db		147	ThrSerTyrProAsnProLysSerPheSerGluPheValGluLysPheMetGlnGlyGln	161
Qy		528	CTGGGTACCGCTCCTGGTTTGTAGCACGTCGCAGAGTTTCTGGGAGCACCCGATGGACTCG	587
Db		167	ValProTyrGlySerTrpTyrAspHisValLysAlaTrpTrpGluLysSerLysAenSer	186
Qy		588	AACGTGCTTTTTCTCAAGTATGAAGACATCATCGGACCCTGGTGACGATGTTGGAGCAG	647
Db		187	ArgValLeuPheMetPheTyrGluAspMetLysGluAspIleArgArgGluValValLys	206
Qy		648	CTGGCCAGAGTTCTCTGGGGGTCTGTGCACAAGGCCCCAGCTGGAAGCCCTGACGGAGCAC	707
Db		207	LeuIleGluPheLeuGluArgLysProSerAlaGluLeuValAspArgIleValGlnHis	226
Qy		708	TGCCAC-----CAGCTCGTGGACCCAG	728
Db		227	ThrSerPheGlnGluMetLysAsnProSerThrAsnTyrThrMetMetProGluGlu	246
Qy		729	TGCTGCAACGCTGAGGCCCTGCC---GTGGCCGGGGAGAAGATTGGGCTGTGGAAGCAG	785
Db		247	MetMetAsnGlnLysValSerProPheMetArgLysGlyIleIleGlyAspTrpLysAsn	266
Qy		786	ATCTTCACCGCTCCCATGAATCAGAAAGTTTGACTTGGTGATAAACAGAAGATGGGAAAG	845
Db		267	HisPheProGluAlaLeuArgGluArgPheAspGluHisIstfLysGlnMetLysAsp	286
Qy		846	TGTGACCTCACGTTT 860	
Db		287	CysThrValLysPhe 291	
RESULT 54				
Q5HZV7 XENTR				
ID	Q5HZV7 XENTR	PRELIMINARY; PRT;	287 AA.	
AC	Q5HZV7:			
DT	15-FEB-2005,	integrated into UniProtKB/TrEMBL.		
DT	15-FEB-2006,	sequence version 11.		
DT	07-FEB-2006,	entry version 11.		
DE	Sulfotransferase family,	cytosolic, 1A, phenol-prefering, member 1.		
GN	Name=sulfat1-prov;			
OS	Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;			
OC	Xenopodinae; Xenopus; Silurana.			
NCBI_TaxID=8364;				
[1]				
RN	NCLEOTIDE SEQUENCE.			
RP	TISSUE=Whole body.			
RC	MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;			
RX	Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,			
RA	Altshul S.F., Jordan B., Buecow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Zeeberg H., Moore T., Max S.I., Wang J., Haieh F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,			
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,			
RA	Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S.C., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fahney J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakesley R., Touchman J.W., Green E.D., Dickson M.C.,			
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,			
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,			


```

Db      198 LysValMetArgPheLeuAspLysAspLeuSerGluGluValLeuGluLysValHis 217
      ::::  |||||  :::  |||||
Qy      696 CTGACGGAGCACTGCCACCACCTGGTGACAGCTGCTCAAGCTGAGGCCCTGCC-- 752
      ||:::  ||||  :::  ||||
Db      218 LeuSerPheAspHisMetLysAspAsnProMetAlaAsnPheSerAlaPheProSer 237
      ::::  ||||  :::  ||||
Qy      753 -----GTGGCCGGGGAAGAGTTGGGCTGTGGGAAG 782
      ::::  ||||  :::  ||||
Db      238 AspValValAspGlnSerGlnTyrLysPheMetArgLysGlyLysValGlyAspTrpLys 257
      ::::  ||||  :::  ||||
Qy      783 GACATCTTCACCGTCTCCATGAATGAGAAGTTGACTTGGTGTATATAACAGAGATGGGA 842
      |||||  |||||  |||||  |||||
Db      258 SerHisPheThrValGlnGlnAsnGluMetPheGluGluLysTyrGlnGlnMetHis 277
      ::::  ||||  :::  ||||
Qy      843 AAGTGTGACCTCAGCTTTGACTTT 866
      :::  |||  :::
Db      278 GlySerAlaMetLysPheArgTyr 285
      :::  |||  :::

RESULT 55
STIE1_MOUSE
ID STIE1_MOUSE STANDARD; PRT; 295 AA.
AC P49891;
DT 01-OCT-1996, integrated into UniProtKB/Swiss-Prot.
DT 15-JUL-1999, sequence version 2.
DT 07-PEB-2006, entry version 47.
DE (Sulfoltransferase, estrogen, testis isoform (EC 2.8.2.4)
DE (Sulfoltransferase, estrogen-prefering).
GN Name=Sult1el; Synonyms=Ste;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA].
RC STRAIN=C57BL/KSJ-DB/DB; TISSUE=Testis;
RX MEDLINE=95269690; PubMed=7750469; DOI=10.1210/en.136.6.2477;
RA Song W.-C., Moore R., McLachlan J.A., Negishi M.;
RT "Molecular characterization of a testis-specific estrogen
RT sulfoltransferase and aberrant liver expression in obese and
RT diabetic C57BL/KsJ-db/db mice.";
RL Endocrinology 136:2477-2484(1995).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS) IN COMPLEX WITH ADENOSINE
RP 3',5'-BISPHOSPHATE (PAP), AND SEQUENCE REVISION TO 113.
RC STRAIN=C57BL/KSJ-DB/DB; TISSUE=Testis;
RX MEDLINE=98020506; PubMed=9360604;
RA Kakuta Y., Pedersen L.G., Carter C.W., Negishi M., Pedersen L.C.;
RT "Crystal structure of estrogen sulfoltransferase.";
RL Nat. Struct. Biol. 4:904-908(1997).
CC -|- FUNCTION: Sulfation of estradiol and estrone. May control the
CC level of the estrogen receptor by sulfolylating free estradiol.
CC -|- CATALYTIC ACTIVITY: 3'-phosphoadenylyl sulfate + estrone =
CC adenosine 3',5'-bisphosphate + estrone 3-sulfate.
CC -|- SUBUNIT: Homodimer (By similarity).
CC -|- SUBCELLULAR LOCATION: Cytoplasm.
CC -|- TISSUE SPECIFICITY: Testis and at very low level in the placenta.
CC -|- MISCELLANEOUS: Abnormal high expression in liver in obese and
CC diabetic C57BL/KSJ-DB/DB strain mice. Female > male. Normal
CC level in liver.
CC -|- SIMILARITY: Belongs to the sulfoltransferase family.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; S78182; AAB34320.1; -; mRNA.
DR PDB; 1AQY; X-ray; A/B=1-295.
DR PDB; 1AQY; X-ray; A/B=1-295.
DR PDB; 1B06; X-ray; A/B=1-295.
DR IntAct; P49891; -.
DR Ensembl; ENSMUSG00000029272; Mus musculus.
DR MGI; MGI:98431; Sult1el.

```

```

DR LinkHub; P49891; -.
DR GO; GO:0008210; P:estrogen metabolism; IMP.
DR GO; GO:0007565; P:pregnancy; IMP.
DR InterPro; IPR000863; S:Sulfoltransferase.
DR Pfam; PF00685; S:Sulfoltransferase_1; 1.
DR ProDom; PD001218; S:Sulfoltransferase; 1.
DR 3D-structure; Lipid-binding; Steroid-binding; Transferase.
KW CHAIN 1 295
FT ESTROGEN SULFOTRANSFERASE, TESTIS
FT isoform.
FT /FTID=PRO_0000085154.
FT PAPS.
FT NP_BIND 48 53
FT NP_BIND 130 138
FT NP_BIND 193 229
FT NP_BIND 257 259
FT ACT_SITE 108 108
FT CONFLICT 113 113
FT HELIX 8 11
FT STRAND 12 15
FT TURN 16 17
FT STRAND 18 21
FT HELIX 22 24
FT TURN 25 27
FT HELIX 28 32
FT TURN 33 33
FT TURN 38 39
FT STRAND 41 45
FT TURN 47 48
FT STRAND 49 50
FT HELIX 51 62
FT TURN 63 64
FT STRAND 75 78
FT HELIX 79 79
FT STRAND 81 81
FT TURN 82 83
FT STRAND 85 86
FT TURN 87 89
FT STRAND 90 91
FT HELIX 92 97
FT TURN 98 98
FT STRAND 101 102
FT STRAND 104 107
FT HELIX 111 113
FT STRAND 114 114
FT HELIX 116 120
FT TURN 121 122
FT STRAND 124 129
FT HELIX 132 145
FT STRAND 146 146
FT TURN 147 148
FT STRAND 149 149
FT STRAND 153 153
FT HELIX 155 164
FT TURN 165 165
FT STRAND 167 167
FT TURN 168 169
FT HELIX 172 182
FT TURN 183 184
FT STRAND 185 185
FT TURN 186 187
FT STRAND 188 192
FT HELIX 193 198
FT TURN 200 210
FT TURN 211 212
FT HELIX 217 226
FT TURN 227 227
FT STRAND 228 228
FT TURN 229 234
FT STRAND 236 242
FT TURN 243 243
FT STRAND 245 247
FT STRAND 248 248
FT TURN 250 252
FT STRAND 253 253

```


FT STRAND 255 256
FT STRAND 261 261
FT HELIX 263 266
FT TURN 267 267
FT HELIX 270 284
FT TURN 285 286
SQ SEQUENCE 295 AA; 35590 MW; 8E85AB47952BFB1C CRC64;

Alignment Scores:

Pred. No.: 1.16e-21 Length: 295
Score: 431.00 Matches: 94
Percent Similarity: 53.0% Conservative: 57
Best Local Similarity: 33.0% Mismatches: 116
Query Match: 9.8% Indels: 18
DB: 1 Gaps: 3

US-10-768-158-1 (1-2419) x ST1E1_MOUSE (1-295)

```
QY 60 GAGTTCGAGAGCAAGTACTTCGAGTTCATGGCGTGGCGTCCGCCCTTCTGCGCGGG 119
Db 7 GluTyrTyrGluValPheGlyGluPheArgGlyValLeuMetAspLysArgPheThrLys 26

QY 120 AAGATGGAGGAGATCCCAACTTCCTCCGCTGGCGCCAGCGCTGTGGATGCTCACTTAC 179
Db 27 TyrTrpGluaspValGluMetPheLeuAlaArgProaspPheValIleAlaThrLys 46

QY 180 CCCAGTCCGCGCACCAGCTTGTGTCAGGAGGTGGTCTACTTGGTGAGCCAGGCGCTGAC 239
Db 47 ProLysSerGlyThrThrTrpLysSerGluValValTyrMetIleTyrLysGluGlyasp 66

QY 240 CCGATGATGATCGGCTTGATGACATCGACGAGCAGCTCCGCGTCTCGGATGCCACAG 299
Db 67 ValGluLysCysLysGluaspAlaIlePheAsnArgIleProTyrLeuGluCysArgAsn 86

QY 300 CCG-----GGCTGACATCATCAGGAACCTGACCTCCCGGCTCATCAAG 347
Db 87 GluaspLeuIleAsnGlyIleLysGlnLeuLysGluLysGluSerProArgIleValLys 106

QY 348 AGCCACCTGCTACCGCTTCTGCGCTCTGACCTCCCAATGGAGACTCCAAGTCTATC 407
Db 107 ThrHisLeuProLysLeuLeuProAlaSerPheTrpGluLysAsnCysLysMetIle 126

QY 408 TATATGGCTCGCAACCCCAAGATCTGGTGTGCTTATATATGATTTCCACGCTCTCTG 467
Db 127 TyrLeuCysArgAsnAlaLysaspValAlaValSerTyrTyrPheLeuLeuMetIle 146

QY 468 CGGACCATGAGCTACCGAGCACCTTCAAGATTTCTCCGAGGTTTATGATGATGAAG 527
Db 147 ThrSerTyrProAsnProLysSerPheSerGluPheValGluLysPheMetGlnGlyGln 166

QY 528 CTGGGCTACGCTCTCTGTTTGGAGCACCTGAGGAGTTCTGGGAGCACCGCATGGACTCG 587
Db 167 ValProTyrGlySerTyrTrpAspHisValLysAlaTrpTrpGluLysSerLysAsnSer 186

QY 588 AACGTGCTTTTCAAGATGATGAACATGATCGGAGCTGGTGGTACGATGGTGGAGCAG 647
Db 187 ArgValLeuPheMetPheTyrGluaspMetLysGluaspIleArgArgGluValValLys 206

QY 648 CTGGCCAGATTCCTGGGGGTCTCTGTGACAAAGCCAGCTGGAGCCCTCACGAGCAG 707
Db 207 LeuIleGluPheLeuGluArgLysProSerAlaGluLeuValAspArgIleGlnHis 226

QY 708 TGCCAC-----CAGCTGGTGAGCAG 728
Db 227 ThrSerPheGlnGluMetLysAsnAsnProSerThrAsnTyrThrMetMetProGluGlu 246

QY 729 TGCTGAACGCTGAGCCCTGCC---GTGGCCGGGGAAGAGTTGGCTGTGGAAGCAG 785
Db 247 MetMetAsnGlnLysValSerProPheMetArgLysGlyIleIleGlyAspTrpLysAsn 266

QY 786 ATCTTCACCGCTCCATGAATGAGAGTTTGCATCTGGTGATTAACAGAAAGTGGAAAG 845
Db 267 HisPheProGluAlaLeuArgGluArgPheAspGluHisTyrLysGlnGlnMetLysAsp 286
```

QY 846 TGTGACCTCACGTTT 860
Db 287 CysThrValLysPhe 291

RESULT 56

Q8JZX7_MOUSE
ID Q8JZX7_MOUSE PRELIMINARY; PRT; 295 AA.
AC Q8JZX7;
DT 01-OCT-2002, integrated into UniProtKB/TrEMBL.
DT 01-OCT-2002, sequence version 1.
DT 07-FEB-2006, entry version 19.
DE Sulfotransferase family 1E, member 1.
GN Name=Sult1e1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Uterus;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN (2)
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Uterus;
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
Distributed under the Creative Commons Attribution-NoDerivs License

DR EMBL; BC034891; AAK34891.1; -; mRNA.
DR HSP; P49891; IAAQ.
DR SMR; Q8JZX7; 7-294.
DR Ensembl; ENSMUSG00000029272; Mus musculus.
DR MGI; MGI:98431; Sult1e1.
DR GO; GO:0008210; P:estrogen metabolism; IMP.
DR GO; GO:0007565; P:pregnancy; IMP.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransfer_1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Transferase.
SQ SEQUENCE 295 AA; 35590 MW; 5C2848FBD63E1F44 CRC64;

Alignment Scores:

Pred. No.: 1.16e-21 Length: 295
Score: 431.00 Matches: 94
Percent Similarity: 53.0% Conservative: 57
Best Local Similarity: 33.0% Mismatches: 116
Query Match: 9.8% Indels: 18
DB: 2 Gaps: 3

```
US-10-768-158-1 (1-2419) x Q8JZX7_MOUSE (1-295)
Qy 60 GAGTTCGAGACAGTACTTCAGTTCCATGGCGTGGCGCTGCCGCCCTTCTGCGCGGG 119
Db 7 GluTyrTyrGluValPheGlyGluPheArgGlyValLeuMetAspLysArgPheThrLys 26
Qy 120 AAGATGGAGGAGATGCCCACTTCCTCGGTGGCGCCAGCGAGTGTGTGATCGTACCTTAC 179
Db 27 TyrTrpGluAspValGluMetPheLeuAlaArgProAspAspLeuValIleAlaThrTyr 46
Qy 180 CCCAAGTCCGACACAGCTTCTGTCAGGAGTGTGTCTACTTGTGTGAGCCAGGCGCTGAC 239
Db 47 ProLysSerGlyThrTyrTrpIleSerGlyValValTyrMetIleTyrLysGluGlyAsp 66
Qy 240 CCCGATGAGTCGGTTCGATGAACATCGACGAGCTCCGGTCTCGGATGATCCACAG 299
Db 67 ValGluLysCysLysGluAspAlaIlePheAsnArgIleProTyrLeuGluCysArgAsn 86
Qy 300 CCG-----GGCTCGACATCATCAAGGAAGTCACTCTCCCGGCTCATCAAG 347
Db 87 GluAspLeuIleAsnGlyIleLysGlnLeuLysGluLysGluSerProArgIleLeuLys 106
Qy 348 AGCCACCTCGCCTACCGCTTTCTGCTCTGACCTCCAAATGGAGACTCCAAAGTCTATC 407
Db 107 ThrHisLeuProProLysLeuLeuProAlaSerPheTrpGluLysAsnCysLysMetIle 126
Qy 408 TATATGGCTCGAACCCCAAGATCGTGTGTCTTATTCATGATTCACCGCTCTCTG 467
Db 127 TyrLeuCysArgAsnAlaLysAspValAlaValSerTyrTyrPheLeuLeuMetIle 146
Qy 468 CGGACCATGAGTACCGGAGGACCTTCAAGAAATCTCCCGAGGTTTATGAATGATAAG 527
Db 147 ThrSerTyrProAsnProLysSerPheSerGluPheValGluLysPheMetGlnGlyGln 166
Qy 528 CTGGCTACGGCTCTCTGTTTGGACAGCTGCGAGGAGTTCTGGGACACCGCATGACTCG 587
Db 167 ValProTyrGlySerTyrTyrAspHisValLysAlaTrpTrpGluLysSerLysAsnSer 186
Qy 588 AACGTCTTTTCTCAAGTATGAACATGATCGTGGGACCTGGTGACATGGTGGAGCAG 647
Db 187 ArgValLeuPheMetPheTyrGluAspMetLysGluAspIleArgArgGluValValLys 206
Qy 648 CTGGCCAGATCTCTGGGGGTCTCTGTCACAAGGCCAGCTGGGAAGCCCTGACGGAGCAG 707
Db 207 LeuIleGluPheLeuGluArgLysProSerAlaGluLeuValAspArgIleValGlnHis 226
Qy 708 TGCCAC-----CAGCTGGTGACCCAG 728
Db 227 ThrSerPheGlnGluMetLysAsnAsnProSerThrAsnTyrThrMetMetProGluGlu 246
Qy 729 TGCTGCAACGCTGAGGCCCTGCC---GTGGCCGGGGAAGAGTTGGCTGTGGAGGAC 785
Db 247 MetMetAsnGlnLysValSerProPheMetArgLysGlyIleIleGlyAspTrpLysAsn 266
Qy 786 ATCTTCACCGCTCCATCAATGAGAAGTTTGACTTGGTGTATAACACAGAAGTGGAAAG 845
Db 267 HisPheProGluAlaLeuArgGluArgPheAspGluHisTyrLysGlnGlnMetLysAsp 286
Qy 846 TGTGACCTCAGCTTT 860
Db 287 CysThrValLysPhe 291
RESULT 57
Q99ND5 RAT
ID Q99ND5 RAT PRELIMINARY; PRT; 295 AA.
AC Q99ND5,
DT 01-JUN-2001, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2001, sequence version 1.
DT 07-FEB-2006, entry version 17.
DE Estrogen sulfotransferase.
GN Name=sfe2;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muroidea; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Astapova I.I., Yakovenko A.R., Shchelkunova T.A., Chernov B.K.,
RA Sverdlova P.S., Smirnov A.N., Rubtsov P.M.;
RA "Cloning and Preliminary Characterization of the Rat Estrogen
RT Sulfotransferase Gene 5'-Region.";
RL Mol. Biol. 33:447-453 (1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Astapova I.I., Yakovenko A.R., Morozov I.A., Smirnov A.N.,
RA Rubtsov P.M.;
RA "Cloning and Structural Characterization of Two Genes Encoding Rat
RT Estrogen Sulfotransferase.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA TISSUE=Liver;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci F., Prange C.,
RA Raha S.A., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield Y.S.N., Kravinsky M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RA TISSUE=Liver;
RA Director MGC Project;
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyright by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; AJ306223; CAC27405.3; -, Genomic DNA.
DR EMBL; AJ306224; CAC27405.3; JOINED; Genomic DNA.
DR EMBL; AJ306225; CAC27405.3; JOINED; Genomic DNA.
DR EMBL; AJ306226; CAC27405.3; JOINED; Genomic DNA.
DR EMBL; AJ306227; CAC27405.3; JOINED; Genomic DNA.
DR EMBL; AJ306228; CAC27405.3; JOINED; Genomic DNA.
DR EMBL; AJ306229; CAC27405.3; JOINED; Genomic DNA.
DR EMBL; BC088157; AAB88157.1; -, mRNA.
DR HSSP; P49891; IAQU.
DR SMR; Q99ND5; 7-294.
DR Ensembl; ENSRNOG0000001957; Rattus norvegicus.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransfer_1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Transferase.
SQ SEQUENCE 295 AA; 35362 MW; 6A470865BA8F5B0E CRC64;
Alignment Scores: 1.37e-21 Length: 295
Pred. No.: 430.00 Matches: 92
Score: 52.7% Conservative: 62
Percent Similarity:
```

Best Local Similarity: 31.5% Mismatches: 106
Query Match: 9.8% Indels: 32
DB: 2 Gaps: 4

US-10-768-158-1 (1-2419) x Q99ND5_RAT (1-295)

```
QY 60 GAGTTCGAGCAGCAAGTACTTCAGTTCATGCGTCCGCGCGCTCTGCGCGCGG 119
DB 7 GluTyrTyrgluValPheGlyAspPheHisGlyValLeuMetAspLysLeuPheThrLys 26
QY 120 AAGATGGAGGAGATCCCAACATTCGCGTCCGCGCGCGCGCGCGCTCTGCGCGG 179
DB 27 TyrTrpGluAspValGluThrPheSerAlaAlaPProAspAspLeuValThrTyr 46
QY 180 CCCAAGTCGCGCAGCAGCTTCTGCGAGGAGTGTCTACTTGGTGGAGCGCGCTGAC 239
DB 47 ProLysSerGlySerThrTrpIleGlyGluIleValAspMetIleTyrLysGluGlyAsp 66
QY 240 CCCGATGAGATCGGCTTGATGAACATCGACGAGCAGCTCCCGCTCTGGATACCCACG 299
DB 67 ValGluLysCysLysGluAspAlaIlePheAsnArgIleProTyrLeuGluCysArgAsn 86
QY 300 CCG-----GGCTGGACATCATCAAGGAAGTGAACCTCTCCCGCTCATCAAG 347
DB 87 GluAspLeuIleAsnGlyIleLysGlnLeuLysGluLysGluSerProArgIleValLys 106
QY 348 AGCCACCTGCGCTACCGCTTTCTGCGCTCTGACCTCCCAATGGAGACTTCAAGTGCATC 407
DB 107 ThrHisLeuProAlaLysLeuLeuProAlaSerPheTrpGluLysAsnCysLysIle 126
QY 408 TATATGGCTCGAACCAGCATCTGGTGTGTCTTATATCATGTTCCACGCTCTCTG 467
DB 127 TyrLeuLysArgAsnAlaLysAspValValSerTyrTyrPhePheLeuIle 146
QY 468 CGGACCATGAGTACCGAGCACCTTCAAGATTTCTCGCGAGTGTATGAATGAAG 527
DB 147 LysSerTyrProAsnProLysSerPheSerGluPheValGluLysPheMetGluGln 166
QY 528 CTGGCTACGCTCTGCTGTTTGGACGACGTGAGGAGTCTGGGAGCACCGCATGACCTCG 587
DB 167 ValProTyrGlySerTyrAspHisValLysSerTyrTrpGluLysSerLysAsnSer 186
QY 588 AACGTGCTTTTCTCAAGTAAAGACATCGCATCGGACCTGGTGGTGAAGTGTGAGCAG 647
DB 187 ArgValLeuPheMetPheTyrGluAspMetLysGluAspIleArgGluValLys 206
QY 648 CTGGCCAGATTCCTGGGGTGTCTGTGACAGGCCAGCTGGAAGCCCTGACGAGCAC 707
DB 207 LeuIleGluPheLeu-----GluArgAspProLeuAlaGluLeu 219
QY 708 TGCCACCAAGCTGGTG-----GACCAGTGTCTGCAAC 737
DB 220 ValAspLysIleIleGlnHisThrSerPheGlnGluMetLysAsnAsnProCysThrAsn 239
QY 738 GCTGAGGCCCTGCCC-----GTGGCGCGGGA 764
DB 240 TyrSerMetLeuProGluThrMetIleAspLeuLysValSerProPheMetArgLysGly 259
QY 765 AGATTTGGCTGTGGAAGACATCTTCCCGTCTCCATGATGACAACTTTCACCTGTG 824
DB 260 lleValGlyAspTrpArgAsnHisPheProGluAlaLeuArgGluArgPheGluHis 279
QY 825 TATAACAGAAAGTGGAAAGTGTGACCTCACGTTT 860
DB 280 TyrGlnArgHisMetLysAspCysProValLysPhe 291
```

RESULT 58

ST1B1 MOUSE

ID ST1B1 MOUSE STANDARD; PRT; 299 AA.

AC Q90WG7: Q8C301: Q922T0;

DT 04-JAN-2005, integrated into UniProtKB/Swiss-Prot.

DT 04-JAN-2005, sequence version 2.

DT 07-FEB-2006, entry version 29.

DE Sulfotransferase family cytosolic 1B member 1 (EC 2.8.2.-)
DE (Sulfotransferase 1B1) (DOPA/tyrosine sulfotransferase).
GN Name=Sulf1b1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), FUNCTION, TISSUE SPECIFICITY,
RP AND DEVELOPMENTAL STAGE
RC STRAIN=C57BL/6 X CBA; TISSUE=Liver;
RX PubMed=9644246;
RA Saeki Y., Sakakibara Y., Araki Y., Yanagisawa K., Suiko M.,
RA Nakajima H., Liu M.-C.;
RT "Molecular cloning, expression, and characterization of a novel mouse
RT liver SUL1B1 sulfotransferase.";
RL J. Biochem. 124:55-64(1998).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORMS 1 AND 2).
RC STRAIN=C57BL/6J; TISSUE=Oviduct;
RX PubMed=16141072; DOI=10.1126/science.1112014;
RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,
RA Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,
RA Davis M.J., Wilming L.G., Aidinis V., Allen J.E.,
RA Ambesi-Impombato A., Apweiler R., Aturaliya R.N., Bailey T.L.,
RA Bansal K.P., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
RA Di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G.,
RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,
RA Georgii-Hemming P., Gingeras T.R., Gojorchi T., Green R.E.,
RA Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
RA Hill D., Humnietek L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
RA Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,
RA Kitano H., Kollas A., Krishnan S.P., Kruger A., Kummerfeld S.K.,
RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
RA Matsuda H., Matsuura S., Miki H., Mignone F., Miyake S., Morris K.,
RA Mortagui-Tabar S., Mulder N., Nakano N., Nakauchi H., Ng P.,
RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavese G., Pesole G.,
RA Petrovski N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
RA Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,
RA Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sesca L., Sheng Y.,
RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
RA Sperling S., Stupka E., Sugita K., Sultana R., Takenaka Y., Taki K.,
RA Tammoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,
RA Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,
RA Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Ouackenbush J.,
RA Wallestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
RA Fukuda S., Kanamori-Kakayama M., Suzuki M., Aoki J., Arakawa T.,
RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
RA Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Nimomiya N.,
RA Nishio T., Okada M., Plessy C., Shibata K., Shiraki T., Suzuki S.,
RA Tagami M., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J.,
RA Hayashizaki Y.;
RT "The transcriptional landscape of the mammalian genome.";
RL Science 309:1559-1563(2005).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).
RC STRAIN=FVB/N; TISSUE=Colon;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield J.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences";
CC Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -!- FUNCTION: Catalyzes the sulfate conjugation of many hormones,
CC neurotransmitters, drugs and xenobiotic compounds. Sulfonation
CC increases the water solubility of most compounds, and therefore
CC their renal excretion, but it can also result in bioactivation to
CC form active metabolites. Sulfates L-DOPA and D-DOPA, tyrosine
CC isomers such as DL-m-tyrosine, dopamine and thyroid hormones.
CC -!- SUBUNIT: Monomer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasm (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q9QMG7-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q9QMG7-2; Sequence=VSP_012509;
CC Note=No experimental confirmation available;
CC -!- TISSUE SPECIFICITY: Liver specific.
CC -!- DEVELOPMENTAL STAGE: Expression was detected at very low level in
CC liver from 1 day-old and then gradually increased to the maximum
CC level at 4 weeks old.
CC -!- SIMILARITY: Belongs to the sulfotransferase family.
CC
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC EMBL; U92076; AAD09249.1; -; mRNA.
CC EMBL; AF022894; AAD01746.1; -; mRNA.
CC EMBL; AK033498; BAC28321.1; -; mRNA.
CC EMBL; AK087595; BAC39939.1; -; mRNA.
CC EMBL; BC024361; AA024361.1; -; mRNA.
CC PIR; J0186; J0186.
CC HSP; P49888; 1HV3.
CC SMR; Q9QMG7; 5-236.
CC EMBL; ENSMUSG0000029269; Mus musculus.
CC MG1; MG1:2136282; Sult1b1.
CC GO; GO:0005829; Cytoplasm; TAS.
CC GO; GO:0004062; Fatty acid sulfotransferase activity; IDA.
CC InterPro; IPR000863; Sulfotransferase.
CC Pfam; PF00685; Sulfotransferase; 1.
CC ProDom; PD001218; Sulfotransferase; 1.
CC Alternative splicing; Lipid metabolism; Steroid metabolism;
CC Sulfotransferase.
CC CHAIN 1 299 Sulfotransferase family cytosolic 1B
CC member 1, 0000085162.
CC /FTID=PRO_0000085162.
CC FT NP_BIND 48 53 PAPS (By similarity).
CC FT NP_BIND 131 139 PAPS (By similarity).
CC FT NP_BIND 194 230 PAPS (By similarity).
CC FT ACT_SITE 109 109 Proton acceptor (By similarity).
CC FT BINDING 259 259 PAPS (By similarity).
CC FT VARSPIC 299 299 A -> ALAFTNFEIIGFSLKFS (in isoform 2).
CC /FTID=VSP_012509.
CC FT CONFLICT 55 55 S -> T (in Ref. 2).
CC FT CONFLICT 137 137 A -> P (in Ref. 1; AAD09249).
CC FT CONFLICT 275 275 E -> K (in Ref. 2).
CC SQ SEQUENCE 299 AA; 34901 MW; BFD0E0909AD482D CRC64;
CC
CC Alignment Scores:
CC Pred. No.: 1.49e-21 Length: 299
CC Score: 429.50 Matches: 93
CC Percent Similarity: 53.7% Conservative: 51
CC Best Local Similarity: 34.7% Mismatches: 97

Query Match: 9.8% Indels: 27
DB: 1 Gaps: 4
US-10-768-158-1 (1-2419) x STB1_MOUSE (1-299)
QY 126 GAGGAGATCGCCCAACTTCCCGCGTCCGCGCCAGCAGCTGGATCGTACCTACCCCAAG 185
DB 29 GluArgileGluGluPheGlnSerThrProGlyAspileValilleThrThrProLys 48
QY 186 TCGCGCACCACTTGGTTCGAGGAGTGTCTACTGTGTGAGCCAGCGGCGTACCCCGAT 245
DB 49 SerGlyThrThrTrpLeuSerGluileValAspMetValLeuAanAspGlyAsnValglu 68
QY 246 GAGATCGCGTGTGATGACATCGACGACGCTCCGCTCTGGAGTACCCACACCGCGGC 305
DB 69 LysCysLysArgAspValilleThrSerLysValProMetLeuGluLeuSerValProgly 88
QY 306 CTG-----GACATCATCAAGAACTGACCTCTCCCGCCCTCATCAAGAGC 350
DB 89 IleArgileSerGlyValGluLeuLeuLysThrProSerProA-gilleLysThr 108
QY 351 CACCTGCGCTACCGCTTCTGCGCTCTGACCTCCACAATGGAGACTCCAAAGTTCATCTAT 410
DB 109 HisLeuProLleAspLeuLeuProLysSerPheTrpGluAenLysCysLysMetIleTyr 128
QY 411 ATGGCTCGCACCCCAAGGATCTGGTGTGTCTATTATCAGTTCCACCGCTCTCTCGG 470
DB 129 LeuAlaArgAenGlyLysAspValAlaValSerTyrTyrHisPheAspLeuMetAsnSer 148
QY 471 ACCATGAGTACCGGAGGACCTTTCAAGAATTTCCGCGAGGTTTATGAATGATAAGCTG 530
DB 149 IleAsnProLeuProGlyThrTrpGluGluTyrLeuGluLysPheLeuAlaGlyAsnVal 168
QY 531 GGCTACGGCTCCTGGTTGAGCAGCTGAGAGTTCGCGAGCACCGGACCGGACTCGAAC 590
DB 169 AlaTyrGlySerTrpPheAspHisValysSerTrpTrpGluLysArgGluGluHisPro 188
QY 591 GTGCTTTTCTCAAGTATGAGACATGCATCGGACCTGGTGACGATGGTGGACGAGCTG 650
DB 189 LeuLeuTyrLeuTyrTyrGluGluLeuLysGlnAsnProLysLysGluilleLysLys 208
QY 651 GCCAGATTCTCTGGGGGTCTCTGTGACAAAGCCCGAGCTGGAAGCCCTGACGAGCAGCTGC 710
DB 209 AlaSerPheLeuAspLysThrLeuAspGluGluAlaLeuAspArgIleValHisThr 228
QY 711 -----CACAGCTGGTGACACGCTGCTGCAACCGCTGCAACCGTGGAGCCCTG 749
DB 229 SerPheGluMetMetLysGluAsnProLeuVal-----AsnTyrThrHisLeu 244
QY 750 CCGCTGGGC-----CGGGAAGAGTTGGGCTG 776
DB 245 ProThrAlaMetMetAspHisSerLysSerProPheMetArgLysGlyileValGlyAsp 264
QY 777 TGAAGAGACATCTTCCACCGTCTCCATGAATGAGAAGTTTGACTTGGTGTATAACAGAG 836
DB 265 TrpLysAsnTyrPheThrMetThrGlnThrGluGlnPheAspAlaValTyrLysLysLys 284
QY 837 ATGGGAAGTGTGACCTCACCTTT 860
DB 285 MetSerGlyThrThrLeuGluPhe 292
RESULT 59
Q91W19_MOUSE PRELIMINARY; PRT; 263 AA.
AC Q91W19; Integrated into UniProtKB/TrEMBL.
DT 01-DEC-2001, sequence version 1.
DT 01-DEC-2001, sequence version 1.
DT 07-FEB-2006, entry version 24.
DE Sulfotransferase family 1A, phenol-preferring, member 1.
GN Name=Sult1a1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;

RL Mol. Biol. (Mosk.) 36:635-642(2002).
 CC -|- FUNCTION: Sulfation of estrone and estradiol. May control the
 CC level of the estrogen receptor by sulfonylating free estradiol.
 CC -|- CATALYTIC ACTIVITY: 3'-phosphoadenylyl sulfate + estrone =
 CC adenosine 3',5'-bisphosphate + estrone 3-sulfate.
 CC -|- SUBUNIT: Homodimer (By similarity).
 CC -|- SUBCELLULAR LOCATION: Cytoplasm.
 CC -|- TISSUE SPECIFICITY: Liver of young mature males and uterus.
 CC -|- DEVELOPMENTAL STAGE: Expressed only in the liver of young adult
 CC animals (100 days old) and is absent in the prepubertal male
 CC days old), senescent male (800 days old) and female liver.
 CC -|- INDUCTION: Induced by androgens and suppressed by estrogens. The
 CC expression is under the influence of pituitary growth hormone and
 CC thyroid hormone. Is regulated by progesterone in the uterus.
 CC -|- SIMILARITY: Belongs to the sulfotransferase family.
 CC
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC
 CC
 DR EMBL; M86758; AA041128.1; -; mRNA.
 DR EMBL; S76489; AB33441.1; -; mRNA.
 DR EMBL; AJ131835; CA010515.2; -; Genomic DNA.
 DR EMBL; AJ298109; CA010515.2; JOINED; Genomic DNA.
 DR EMBL; AJ298110; CA010515.2; JOINED; Genomic DNA.
 DR EMBL; AJ298111; CA010515.2; JOINED; Genomic DNA.
 DR EMBL; AJ298112; CA010515.2; JOINED; Genomic DNA.
 DR EMBL; AJ298113; CA010515.2; JOINED; Genomic DNA.
 DR EMBL; AJ298114; CA010515.2; JOINED; Genomic DNA.
 DR PIR; A41930; A41930.
 DR HSRP; P49891; LAOU.
 DR SMR; P49889; 7-294.
 DR Ensembl; ENSRNOG00000001957; Rattus norvegicus.
 DR RGD; 3776; See.
 DR InterPro; IPR000863; Sulfotransferase.
 DR Pfam; PF00685; Sulfotransferase 1.
 DR ProDom; PD001218; Sulfotransferase; 1.
 KW Direct protein sequencing; Lipid-binding; Steroid-binding;
 KW Transferase.
 FT CHAIN 1 295 Estrogen sulfotransferase, isoform 3.
 FT /FTID=PRO_0000085157.
 FT NP_BIND 48 53 PAPS (By similarity).
 FT NP_BIND 130 138 PAPS (By similarity).
 FT NP_BIND 193 229 PAPS (By similarity).
 FT NP_BIND 257 259 PAPS (By similarity).
 FT ACT_SITE 108 108 Proton acceptor (By similarity).
 FT CONFLICT 150 150 P -> Q (in Ref. 2).
 FT CONFLICT 238 238 T -> I (in Ref. 2).
 FT CONFLICT 295 295 L -> P (in Ref. 2).
 SQ SEQUENCE 295 AA; 35416 MW; A77807A21DD2E7EB CRC64;
 Alignment Scores:
 Pred. No.: 1.61e-21 Length: 295
 Score: 429.00 Matches: 92
 Percent Similarity: 53.3% Conservative: 60
 Best Local Similarity: 32.3% Mismatches: 115
 Query Match: 9.7% Indels: 18
 Gaps: 3
 DB: 1
 US-10-768-158-1 (1-2419) x ST1E3_RAT (1-295)
 QY 60 GAGTTCGAGAGCAAGTACTTCGAGTTCCATGCGTGGCGTCCGCCCTTCTGCGCGGG 119
 Db 7 GlutylTyrAspValPheGlyAspPheHisGlyPheLeuMetAspLysArgPheThrLys 26
 QY 120 AAGATGGAGGAGATCCCAACTTCCCGTGGCGGCCCGACGAGTGTGGATCGTACCTAC 179
 Db 27 TyrTrpGluAspValGluThrPheLeuAlaArgProAspAspLeuLeuValThrTyr 46
 QY 180 CCCAAGTCCGACACAGCTTCTGACGAGGTGGTGTCTACTTGGTGGACGGCGGCTGAC 239
 Db 47 ProLysSerGlySerThrTrpIleValAspMetIleTyrLysGluGlyAsp 66
 QY 240 CCCGATGAGATCGGGTTCATGAACATCGACGAGCAGCTCCCGTCTCTGGAGTACCCACAG 299

Db 67 ValGluLysCysLysGluAspAlaLeuPheAsnArgIleProAspLeuGluCysArgAsn 86
 QY 300 CCG-----GGCTTGGACATCATCAAGAACTACCTCTCCCGCCTCATCAAG 347
 Db 87 GluAspLeuIleAsnGlyIleLysGlnLysGluLysGluSerProArgIleValLys 106
 QY 348 AGCCACCTGCGCTACCGCTTCTGCTCTGACCTCCACATGGAGACTCCAAGTCTATC 407
 Db 107 ThrHisLeuProAlaLysLeuLeuProAlaSerPheTrpGluLysAsnCysLysIleIle 126
 QY 408 TATATGGCTCCCAACCCCAAGGATCTGGTGTGTCTTATTATTCAGTTCCACCGCTCTCTG 467
 Db 127 TyrLeuCysArgAsnAlaLysAspValValSerTyrTyrPhePheLeuIleMet 146
 QY 468 CGGACCATGACTACCGAGGACCTTTCAAGAAATCTGCCGAGGTTTATGAATGATAAG 527
 Db 147 LysSerTyrProAsnProLysSerPheSerGluPheValGluLysPheMetGluGlyGln 166
 QY 528 CTGGGCTACCGCTCTGCTTGGACGAGTCTGGGAGCACCACCGATCGACTCG 587
 Db 167 ValProTyrGlySerTyrAspHisValLysSerTrpTrpGluLysSerLysAsnSer 186
 QY 588 AACCTGCTTTTCTCAAGATGAAGACATCGATCGGACCTGCTGTCAGCATGGTGGAGCAG 647
 Db 187 ArgValLeuPheMetPheTyrGluAspMetLysGluAspIleArgGluValValLys 206
 QY 648 CTGCCACGATTCCTGGGGGTCTCTGTGACAAAGCCCGAGCTGGAAAGCCTGACCGAGCAC 707
 Db 207 LeuIleGluPheLeuGluArgAspProSerAlaGluLeuValAspArgIleGlnHis 226
 QY 708 TGC-----CACCAGCTGGTGGACGAGTCTGCAACGCTGAGGCGCCCTGCC----- 752
 Db 227 ThrSerPheGlnGluMetLysAsnAsnProCysThrAsnTyrSerMetLeuProGluThr 246
 QY 753 -----GTGGCGCGGGAAGAGTGTGGCTGTGGAAGGAC 785
 Db 247 MetIleAspLeuLysValSerProPheMetArgLysGlyIleValGlyAspTrpLysAsn 266
 QY 786 ATCTTCCCGCTCCATGAATGAGAACTTGGTGTGATATAAACAAGAGATGGGAAG 845
 Db 267 HisPheProGluAlaLeuArgGluArgPheGluHisTyrGlnGlnMetLysAsp 286
 QY 846 TGTGACCTCAGCTTT 860
 Db 287 CysProValLysPhe 291
 RESULT 61
 Q3T0S9 BOVIN PRELIMINARY; PRT; 295 AA.
 AC Q3T0S9;
 DT 11-OCT-2005, integrated into UniProtKB/TrEMBL.
 DT 07-MAR-2006, entry version 4.
 DE Hypothetical protein.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
 OC Pecora; Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=Crossbred x Angus; TISSUE=ileum;
 RA Moore S., Alexander L., Brownstein M., Guan L., Lobo S., Meng Y.,
 RA Tanaguchi M., Wang Z., Yu J., Prange C., Schreiber K., Shennen C.,
 RA Wagner L., Bala M., Barbazuk S., Barber S., Babakoff R., Beland J.,
 RA Chun E., Del Rio L., Gibson S., Hanson R., Kirkpatrick R., Liu J.,
 RA Matsuo C., Mayo M., Santos R.R., Stott J., Tsai M., Wong D.,
 RA Siddiqui A., Holt R., Jones S.J., Marra M.A.;
 RL Submitted (AUG-2005) to the EMBL/GenBank/DBJ databases.
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License

[illegible]


```

QY 402 GTCATCTATATGGCTCGCAACCCCAAGGATCTGGTGGTGCTTATTATCAGTCCACGCG 461
Db   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 137 llettyrvalaAaGAsAlaLysAspAenLeuValSerTyrTrpHisPheHisArg 156
Db   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 462 TCTCTCGCGACCATGAGCTACCGAGCCACCTTTCAAGAAATTCGCGGAGGTTTATGAAT 521
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 157 MetAenLysValLeuProAspProGlyThrIleGluGluPheThrGluLysPheMetAen 176
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 522 GATAGCTGGCTACGGCTCTCTGGTTGAGCAGCTGCAGGAGTTCTGGGAGCACCGCATG 581
Db   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 177 GlyGluValLeuTrpGlySerTrpTyrAspHisValLysGlyTrpTrpLysAlaLysAsp 196
Db   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 582 GACTCGAACGCTTTTCTCAAGTATGAAGACATGCATCGGACCTGGTGACGATGGTG 641
Db   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 197 LysHisArgIleLeuTyrLeuPheTyrGluAspMetLysGluAenProLysArgGluIle 216
Db   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 642 GAGCAGCTGGCCAGTCTCTGGGGGTCTCTGTGACAAAGGCCAGCTG- 689
Db   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 217 GlnLysIleMetLysPheLeuGluLysAspLeuAspGluGluValLeuAenLysIle 236
Db   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 690 -----GAAGCCCTGACGGAGCATGCCACAGCTG 719
Db   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 237 TyrAenThrSerPheGluIleMetLysAspAenProMetThrAsnTyrThrLysAspPhe 256
Db   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 720 GTGGACCACTGCTGCAACGCTGAGCCCTGCCGCTGGCGCGGGAAGATTGGGCTGGG 779
Db   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 257 ValGlyValMetAenHisSerValSerProPheMetArgLysGlySerValGlyAspTrp 276
Db   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 780 AAGCACTCTTCACCGCTCCCATGAATGAGAAGTTTGATCTGGTGATAAACAAGAATG 839
Db   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 277 LysAenTyrPheThrValAlaLeuAenLysLysPheAspGlnAenTyrLysLysLysMet 296
Db   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 840 GGAAGTGTGACCTCAGCTTT 860
Db   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 297 AlaAspThrSerLeuValPhe 303
Db   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 63
ST1S3_BRARE
ID ST1S3_BRARE STANDARD; PRT; 301 AA.
AC Q772V2;
DT 31-AUG-2004, integrated into UniProtKB/Swiss-Prot.
DT 01-OCT-2003, sequence version 1.
DT 07-FEB-2006, entry version 19.
DE Cytosolic sulfotransferase 3 (EC 2.8.2.-) (SULT1 ST3).
GN Name=sult1st3;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA], ACTIVITY, TISSUE SPECIFICITY, AND ENZYME
RP REGULATION.
RX MEDLINE=22630056; PubMed=12745256; DOI=10.1016/S0003-9861(03)00172-3;
RA Sugahara T., Liu C.-C., Carter G., Pai T.G., Liu M.-C.;
RT "cDNA cloning, expression, and functional characterization of a
RT zebrafish SULT1 cytosolic sulfotransferase.";
RL Arch. Biochem. Biophys. 414:67-73(2003).
CC -!- FUNCTION: Catalyzes the conjugation of sulfate to a variety of
CC xenobiotic and endogenous compounds, including dopamine, T3
CC (3,5,3',5'-tetraiodo-L-thyronine), T4 (thyroxine), estrone, DHEA
CC (dehydroepiandrosterone), flavonoids, isoflavonoids and other
CC phenolic compounds.
CC -!- ENZYME REGULATION: Inhibited by Hg(2+), Co(2+), Zn(2+), Cd(2+),
CC Cu(2+) and Pb(2+) ions. Activated slightly by Mn(2+), Ca(2+) and
CC Mg(2+) ions.
CC -!- BIOPHYSICOCHEMICAL PROPERTIES:
CC pH dependence:
CC Optimum pH is about 5.0 with n-propyl gallate as substrate, and
CC another smaller pH optimum is observed spanning pH 9.5-10.5.
CC Optimum pH is 10.5 with dopamine as substrate;
CC Temperature dependence:
CC Active from 20 to 43 degrees Celsius;

```

```

CC -!- SUBCELLULAR LOCATION: Cytoplasm.
CC -!- SIMILARITY: Belongs to the sulfotransferase family.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC -----
CC EMBL: AY196985; AAP55637.1; -; mRNA.
CC Ensembl: ENSDARG00000018361; Danio rerio.
CC ZFIN: ZDB-GENE-030804-28; sult1st3.
CC GO: GO:0005737; C:cytoplasm; NAS.
CC GO: GO:0008146; F:sulfotransferase activity; IDA.
CC GO: GO:0006805; P:xenobiotic metabolism; IDA.
CC InterPro: IPR000863; Sulfotransferase.
CC Pfam: PF00685; Sulfotransferase_1; 1.
CC ProDom: PD001218; Sulfotransferase; 1.
KW Catecholamine metabolism; Lipid metabolism; Steroid metabolism;
KW Transferase.
KW CHAIN 1 301 Cytosolic sulfotransferase 3.
FT NP_BIND 53 58 PAPS (By similarity).
FT NP_BIND 137 145 PAPS (By similarity).
FT NP_BIND 201 237 PAPS (By similarity).
FT NP_BIND 263 265 PAPS (By similarity).
FT ACT_SITE 115 115 Proton acceptor (By similarity).
SQ SEQUENCE 301 AA; 35377 MW; 2E4D05DA38EB741 CRC64;

Alignment Scores:
Pred. NO.: 2,07e-21 Length: 301
Score: 427.50 Matches: 97
Percent Similarity: 54.1% Conservative: 54
Best Local Similarity: 34.8% Mismatches: 109
Query Match: 9.7% Indels: 19
DB: 1 Gaps: 5

US-10-768-158-1 (1-2419) x ST1S3_BRARE (1-301)
QY 81 GAGTTCCATGGCGGTGCGGCTGCGCCCTTCTGCGCGGGAAGATGGAGGATCGCCAAC 140
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 19 AspPheGluGlyIleSerMetIleHisTyrPheThrAspAenTrpGluLysValLysAen 38
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 141 TTCCCGGTGGCGGCCAGCAGCTGTGGATCGTACCTACCCCAAGTCGGGACGAGCTG 200
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 39 PheGlnAlaArgProAspAspIleLeuIleAlaThrTyrProLysAlaGlyThrThrTrp 58
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 201 CTGACGAGGTGTCTACTTGGTGAGCGAGGCGCTGAC---CCGATGAGATCGGCTG 257
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 59 ValSerTyrIleLeuAspLeuLeuTyrPheGlyAenGluSerProGluArgGlnThrSer 78
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 258 ATGAACATCGACGAGCAGCTCCCGCTCTGGAGTACCCACAGCGCGGCTGGACATCATC 317
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 79 GlnProIleTyrMetArgValProPheLeuGluAlaCysPheGluGlyIleProPheGly 98
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 318 AAGGAACGTG-----ACCTCTCCCGCCTCATCAAGAGCCACTGCCCTAC 362
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 99 ThrGluLeuAlaAspAenLeuProThrSerProArgLeuIleLysThrHisLeuProVal 118
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 363 CGCTTTCTGCCCTCTGACCTCCACAAATGGAGACTCCCAAGTCACTATATGGCTCGCAAC 422
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 119 GlnLeuValProLysSerPheThrGluGlnAenSerLysValValTyrValAlaArgAen 138
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 423 CCCAAGGATCTGGTGGTGCTTATTATACAGTTCCACCGCTCTCTGCGGACCATGAGCTAC 482
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 139 AlaLysAspAenAlaValSerTyrPheHisPheAspArgMetAenMetGlyGlnProGlu 158
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 483 CGAGGCACTTTCAAGAAATTCGCGGAGGTTTATGAATGATAAGCTGGGCTACGGCTCC 542
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 159 ProGlyAspTrpAenThrPheLeuGlnLysPheMetGluGlyArgAenValPheGlyPro 178
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 543 TGGTTTGACGACGTGGAGGAGTCTGGAGCACCAGCATGGAC---TCGAAGTCTTTT 599
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 179 TrpTyrAspHisValAenGlyTyrTrpLysLysLysGlnThrTyrSerAsnIleLeuTyr 198
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 600 CTCAAAGTATGAACATGCATCGGCACTGGTACGATGGTGGAGCAGCTGGCCAGATTC 659

```

Db 199 MetPheTyGluAspMetValGluAenThrThyArgGluValGluArgLeuCysserPhe 218
Qy 660 CTGGGGGTCTCTGTGACAAAGCCAGCTGGAAGCCCTGACGGAGCACTGCCAC----- 713
Db 219 LeuGlyLeuSerThrSerAlaAlaGluArgGluArgIleThrLysGlyValGlnPheAsp 238
Qy 714 ---CAGCTGGTGACAGCTGTCMAAGCTGAGGCCCTGCCCGTG----- 755
Db 239 AlaMetLysGlnAenLysMetThrAsnTySerThrIleProValMetAspPheLysIle 258
Qy 756 -----GGCCGGGAGAGATTGGGCTGTGAAGAGACATCTTACCGCTCCCATG 803
Db 259 SerProPheMetArgLysGlyValGlyAspTrpArgAsnHisPheThrValAlaGln 278
Qy 804 AATGAGAAGTTTGACTGTGTATATAACAGAGATGGGAAAGTGTGACCTCACGTTT 860
Db 279 AenGluGlnPheAspGluValTyLysGlnLysMetLysAenThrThrValLysPhe 297

RESULT 64

ST2S2_BRARE
ID ST2S2_BRARE STANDARD; PRT; 301 AA.
AC Q7ZUS4; Q7T1C7;
DT 31-AUG-2004, integrated into UniProtKB/Swiss-Prot.
DT 01-JUN-2003, sequence version 1.
DT 07-FEB-2006, entry version 2. (EC 2.8.2.-) (SULT1 S72).
DE Cytosolic sulfotransferase 21 (EC 2.8.2.-) (SULT1 S72).
GN Name=sult1st2;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA], ACTIVITY, TISSUE SPECIFICITY, AND ENZYME
REGULATION.
RX MEDLINE=22640919; PubMed=12755695;
RA DOI=10.1046/j.1432-1033.2003.03608.x;
RA Sugahara T., Liu C.-C., Pai T.G., Collodi P., Suiko M., Sakakibara Y.,
RA Nishiyama K., Liu M.-C.;
RT "Sulfation of hydroxychlorobiphenyls. Molecular cloning, expression,
RT and functional characterization of zebrafish SULT1
RL Eur. J. Biochem. 270:2404-2411(2003).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RG NIH - Zebrafish Gene Collection (ZGC) project;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Catalyzes the conjugation of sulfate to a variety of
CC xenobiotic and endogenous compounds, including 2-naphthol,
CC hydroxychlorobiphenyls, T3 (triiodo-L-thyronine), T4 (thyroxine),
CC estrone and DOPA.
CC -!- ENZYME REGULATION: Inhibited by Co(2+), Zn(2+), Cd(2+) and Pb(2+)
CC ions. Inactivated by Hg(2+) and Cu(2+) ions.
CC -!- BIOPHYSICOCHEMICAL PROPERTIES:
CC Optimum pH is 4.75 and 10.5. These two pH optima may correspond
CC to two distinct conformational states of the enzyme;
CC Temperature dependence:
CC Active from 20 to 43 degrees Celsius;
CC -!- SUBCELLULAR LOCATION: Cytoplasm.
CC -!- TISSUE SPECIFICITY: Expressed in liver.
CC -!- SIMILARITY: Belongs to the sulfotransferase family.
CC
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC
CC EMBL; AY181065; AA064984.1; -; mRNA.
CC EMBL; BC047850; AAH47850.1; -; mRNA.
CC HSSP; P49888; 1HY3.
CC Ensembl; ENSDARG00000018361; Danio rerio.
CC ZFIN; ZDB-GENE-030804-27; sult1st2.
CC GO; GO:0005737; C:cytoplasm; NAS.

DR GO; GO:0008146; F:sulfotransferase activity; IDA.
DR GO; GO:0006805; P:xenobiotic metabolism; IDA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase 1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Catecholamine metabolism; Lipid metabolism; Steroid metabolism;
KW Transferase.
FT CHAIN 1 301 Cytosolic sulfotransferase 2.
FT NP_BIND 53 58 PAPS (By similarity).
FT NP_BIND 137 145 PAPS (By similarity).
FT NP_BIND 201 237 PAPS (By similarity).
FT NP_BIND 263 265 PAPS (By similarity).
FT ACT_SITE 115 115 Proton acceptor (By similarity).
FT CONFLICT 111 111 P -> L (in Ref. 1).
FT CONFLICT 226 226 T -> A (in Ref. 1).
FT CONFLICT 239 239 V -> A (in Ref. 1).
FT CONFLICT 285 285 V -> D (in Ref. 1).
SQ SEQUENCE 301 AA; 35364 MW; 5943B7C43E33C621C CRC64;
Alignment Scores:
Pred. No.: 2.07e-21 Length: 301
Score: 427.50 Matches: 99
Percent Similarity: 54.1% Conservative: 52
Best Local Similarity: 35.5% Mismatches: 109
Query Match: 9.7% Indels: 19
DB: 1 Gaps: 5
US-10-768-158-1 (1-2419) x ST2S2_BRARE (1-301)
Qy 81 GAGTTCATGTCGGTGGCTCGCGCTTCGCGCGGGAAGATCGAGAGATCCGCCAAC 140
Db 19 AspPheGluGlyValSerMetThrArgTyPheThrAspAsnTrpGluLysValLysAsn 38
Qy 141 TTCGGGTGGCGCCCGAGCGTGTGGATCTCACTACCCCACTCGCGGACAGCTG 200
Db 39 PheGlnAlaArgProAspIleLeuIleAlaThrTyProLysAlaGlyThrTrp 58
Qy 201 CTGAGGAGGTGTCTACTTGTGTAGCGCGCGCTGAC---CCGATGAGATCGGCTG 257
Db 59 ValSerTyIleLeuAspLeuLeuTyPheGlyAsnGluSerProGluArgGlnThrSer 78
Qy 258 ATGAACATCGACGACGAGCTCCGGTCTCTGGATGACCCACAGCCGCGGCTGACATC 317
Db 79 GlnProIleTyMetArgValProPheLeuGluMetCysPheGlnGlyLeuProLeuGly 98
Qy 318 AAGAACTG-----ACCTCTCCCGCTCATCAAGAGCTCATATATGCTCGCAAC 362
Db 99 ThrGluLeuAlaAspThrLeuProThrSerProArgProIleLeuHisLeuProVal 118
Qy 363 CGCTTTCCTGCTGACCTCCACATGAGACTCCAAAGTCTCATATATGCTCGCAAC 422
Db 119 GlnLeuValProLysSerPheTrpGluGlnAsnSerLysValValTyValAlaArgAsn 138
Qy 423 CCCAAGGATCTGGTGTCTTATATATCAGTTCACCGCTCTCTGCGGACCATGAGCTAC 482
Db 139 AlalysAspAsnAlaValSerTyPheHisPheAspArgMetAsnMetGlyGlnProGlu 158
Qy 483 CGAGGCACCTTTCAGAAATTCGCGGAGGTTTATGAATGATAGTGGCTAGGCTCC 542
Db 159 ProGlyAspTrpAsnThrPheLeuGlnLysPheMetAspGlyArgAsnValPheGlyPro 178
Qy 543 TGGTTTGAGCAGCTGCAGAGGAGTTCGGGAGCACCCTGAGAC---TCGAAGTGTGTTT 599
Db 179 TrpTyAspHisValAsnGlyTyTrpLysLysGlnThrTySerAsnIleLeuTy 198
Qy 600 CTCAGGTATGAACATGATCGGACCTGTGTGACGATGTTGGAGAGCTCGCCAGATT 659
Db 199 MetPheTyGluAspMetValGluAspThrGlyArgGluValAlaAlaLeuLysCysSerPhe 218
Qy 660 CTGGGGGTGCTCTGTGACAGGCCCTGAGGACCTGAGGAGACTGCCACAGCTG 719
Db 219 LeuGlyLeuSerThrSerAlaThrGluArgGluArgIleThrLysGlyValGlnPheAsp 238

QY 720 GTG-----GACCACTGCTGCAACCTGAGGCCCTGCCGCTG----- 755
 Db 239 ValMetLysGlnAsnLysMetThrAsnTyrSerThrLeuProValMetAspPheLysIle 258
 QY 756 -----GGCCGGGGAAGATTGGCTGTGGAAAGGACATCTTCCACCGTCTCCATG 803
 Db 259 SerProPheMetArgLysGlyValGlyAspTrpArgAsnHisPheThrValAlaGln 278
 QY 804 AATGAGAAGTTGACTTGGTGATATAACACAGAGATGGGAAAGTGATGACCTCACGTTT 860
 Db 279 AsnGluGlnPheAspGluValTyrLysGlnLysMetLysAsnThrThrValLysPhe 297

RESULT 65
 Q4FZP1_XENLA
 ID Q4FZP1_XENLA PRELIMINARY; PRT; 302 AA.
 AC Q4FZP1;
 DT 30-AUG-2005, integrated into UniProtKB/TrEMBL.
 DT 30-AUG-2005, sequence version 1.
 DT 07-FEB-2006, entry version 2.
 DE Hypothetical protein.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
 OC Xenopodinae; Xenopus; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Whole;
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.;
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 RT initiative.";
 RL Dev. Dyn. 225:384-391(2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Whole;
 RX MEDLINE=2238257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.F., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.M., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Pahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blackley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Whole;
 RA Klein S., Gerhard D.S.;
 RL submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC -----
 DR EMBL; BC099307; AAH99307.1; -; mRNA.
 DR GO; GO:0008146; F:sulfotransferase activity; IEA.
 DR InterPro; IPR000863; Sulfotransferase.
 DR Pfam; PF00685; Sulfotransfer_1; 1.
 DR ProDom; PD001218; Sulfotransferase; 1.

KW Hypothetical protein.
 SQ SEQUENCE 302 AA; 35792 MW; AF91BD3FB8BF9C9 CRC64;
 Alignment Scores:
 Pred. No.: 2,07e-21 Length: 302
 Score: 427.50 Matches: 89
 Percent Similarity: 53.2% Conservative: 61
 Best Local Similarity: 31.6% Mismatches: 115
 Query Match: 9.7% Indels: 17
 DB: 2 Gaps: 3

US-10-768-158-1 (1-2419) x Q4FZP1_XENLA (1-302)

QY 72 AAGTACTTCGAGTTCATGGCTGGCTGCCGCTTCTGCTCCGCGGGAAGATGGAGAG 131
 Db 4 LysLeuGlnSerMetGluGlyIleArgIleAlaGlyValIleAlaThrAsnTrpGlnGln 23
 QY 132 ATCCCAACTTCCCGGTGCCGCCAGCGAGCTGTGGATCGTCACTACCCCAACATCCGCGC 191
 Db 24 IleArgThrPheGlnAlaArgLeuGlyAspValLeuIleAlaThrTyrProLysSerGly 43
 QY 192 ACCAGTTGCTGAGGAGGTGCTTACTTGTGTAGCCAGGCGGCTGACCCCGATGAGATC 251
 Db 44 ThrThrTrpValGlnGluIleValAspLeuIleLeuAsnGluGlyAsnGluGluIleCys 63
 QY 252 GGCTTGATGAACATCGAGAGCAGCTCCG-----GTCCTGGAGTAGTCCACACGCG 302
 Db 64 ArgArgSerProThrHisGluArgIleProPheValGluLeuHisLeuMetLysPro 83
 QY 303 GGCCTGGACATCATCAAGAACTGACCTCTCCCGCCTCATCAAGAGCCACCTGCCCTAC 362
 Db 84 GlyProGluGluValAsnAlaMetProSerProArgValLeuLysThrHisLeuProVal 103
 QY 363 CGCTTTCGCTCTGACCTCCACAAATGGAGACTCCCAAGGTGATCATATGATGGCTGCGAAC 422
 Db 104 GlnLeuValProProLeuPheTrpLysTyrLysCysLysValIleTyrValAlaArgAsn 123
 QY 423 CCCAAGGATCTGGTGTGCTTATTATCAGTTCCACCGCTCTCTCGGACCATGAGCTAC 482
 Db 124 ProArgAspThrLeuThrSerTyrPheHisPheAspHisMetValLysIleHisProAsp 143
 QY 483 CGAGGCACCTTCAAGAAATCTGCGGAGGTATTATGAATGATAGCTGGGCTACCGCTCC 542
 Db 144 ProGluSerTrpGluGluTyrLeuHisArgPheMetLysGlyAspValGlyTrpGlySer 163
 QY 543 TGGTTTGAGCAGCTGAGGAGTCTGGGAGCACCGCATGAGCTCGAAGCTGCTTTTCTC 602
 Db 164 TrpTyrAspHisValLysGlyPheTrpGluGlnLysAspGluHisAsnIleLeuTyrLeu 183
 QY 603 AAGTATGAAGACATGCTCGGACCTGTCAGCATGTCGAGGAGCTGGGACAGCTGGCCAGATTCTCTG 662
 Db 184 PhePheGluAspIleLysArgAsnSerIleAsnGluIleArgLysValMetArgPheLeu 203
 QY 663 GGGGTGCTCTGTGCAAGGCCACGCTGGAAGCC-----CTGACGGAGCATCTGCCAC 713
 Db 204 AspLysAspLeuSerGluGluValLeuGluLysIleAlaHisLeuSerSerPheAsnGln 223
 QY 714 CAGCTGTGGACAGTCTGCAACGCTGACGCTGAGGCCCTGCC----- 752
 Db 224 MetMetGluAsnProMetAlaAsnTyrSerAlaPheProSerAspValLeuAspGlnSer 243
 QY 753 -----GTGGGCGGGGAAGATTGGGCTGTGGAGGACATCTTCCACCGTCTCC 800
 Db 244 GlnTyrLysPheMetArgLysGlyValGlyAspTrpLysAsnHisPheThrValGln 263
 QY 801 ATGAATGAGAAGTTTGAATGGTGTATATAACAGAAAGATGGGAAAGTGTGACCTCACGTTT 860
 Db 264 GlnAsnGluMetPheGluAlaTyrTyrGlnGlnGlnMetHisGlyCysThrMetLysPhe 283
 QY 861 GACTTT 866
 Db 284 ArgTyr 285

RESULT 66
ID Q2XV97_MACFA PRELIMINARY; PRT; 294 AA.
AC Q2XV97;
DT 20-DEC-2005, integrated into UniProtKB/TrEMBL.
DT 20-DEC-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Sulfotransferase family 1b, estrogen-preferring, member 1.
GN Name=SULT1E1;
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC Cercopithecoidea; Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Liu H., Larbie F., Luu-The V.;
RL Submitted (SEP-2005) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL: DQ228171; ABB76812.1; -: mRNA.
DR GO: GO:0016740; F:transferase activity; IEA.
KW Transferase.
SQ SEQUENCE 294 AA; 35159 MW; B9997AFDBCC353C0 CRC64;

Alignment Scores:
Pred. No.: 4,28e-21 Length: 294
Score: 423.00 Matches: 91
Percent Similarity: 52.3% Conservative: 58
Best Local Similarity: 31.9% Mismatches: 118
Query Match: 9.6% Indels: 18
DB: 2 Gaps: 3

US-10-768-158-1 (1-2419) x Q2XV97_MACFA (1-294)
QY 60 GAGTTCGAGAGCAAGTACTTCGAGTTCATGGCGTCCGCTGCGCCCTTCTGCGCGGG 119
DB 6 AspTyrTyrGluAsnPheGluGluLeuHieGluValLeuMetTyrLysAspPheValLys 25
QY 120 AAGATGGAGAGATGCCAACTTCGCGTCCGCGCCAGGACGTGGTGGATCGTCACTAC 179
DB 26 TyrTrpAsnAspValGluThrPheGlnAlaArgProAspAspLeuValIleAlaThrTyr 45
QY 180 CCAAGTCCGCGACCACTTCGTCGAGGAGTGTCTACTGTGTAGCCAGCGCGCTGAC 239
DB 46 ProLysSerGlyThrTrpValSerGluIleAlaTyrMetIleTyrLysGluGlyAsp 65
QY 240 CCGATGAGATCGGCTTGATGAATCGACGACGCTCCGCTCTCGAGTACCCACAG 299
DB 66 ValGluLysCysLysGluAspValIlePheAsnArgIleProPheLeuGluCysArgLys 85
QY 300 CCG-----GGCTGGACATCATCAAGAACTGACCTCTCCCGCTCATCAAG 347
DB 86 GluAspLeuMetAsnGlyValLysGlnLeuAspGluMetAsnSerProArgIleValLys 105
QY 348 AGCAGCTGCGCTACCGCTTCTGCGCTCTGACCTCCACAAAGGAGACTCCAGGTGATC 407
DB 106 ThrHisLeuProGluLeuLeuProAlaSerPheTrpGluLysAsnCysLysIleIle 125
QY 408 TATATGCTCGCAACCCCAAGGATGGTGTCTTATTATTCAGTTCCACCGCTCTCTG 467
DB 126 TyrLeuCysArgAsnAlaLysAspValAlaValSerPheTyrTyrPhePheLeuMetVal 145
QY 468 CGGACCATGAGTACCGAGCACCTTCAAGAAATCTGCGCGAGGTTTATGATGATAAG 527
DB 146 AlaGlyHisProAsnProGlySerPheProGluPheValGluLysPheMetGlnGlyGln 165
QY 528 CTGGGCTACGGCTCTGGTTTGGACAGGTGACGAGTCTTGGGAGCACCGCATGCACTCG 587
DB 166 ValProTyrGlySerTrpTyrLysHisValLysSerTrpTrpGluLysGluLysSerPro 185

RESULT 66
 Q2XV97_MACFA PRELIMINARY; PRT; 294 AA.
 AC Q2XV97;
 DT 20-DEC-2005, integrated into UniProtKB/TrEMBL.
 DT 20-DEC-2005, sequence version 1.
 DT 07-FEB-2006, entry version 3.
 DE Sulfotransferase family 1b, estrogen-preferring, member 1.
 GN Name=SULT1B1;
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 OC Cercopithecoidea; Cercopithecinae; Macaca.
 ON NCBI_TaxID=9541;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Liu H., Larbie F., Luu-The V.;
 RL Submitted (SEP-2005) to the EMBL/GenBank/DBJ databases.
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC -----
 DR EMBL: DQ228171; ABB76812.1; -; mRNA.
 DR GO: GO:0016740; F:transferase activity; IEA.
 KW Transferase.
 SQ SEQUENCE 294 AA; 35159 MW; B9997AFDBCC353C0 CRC64;
 Alignment Scores:
 Pred. No.: 4,28e-21 Length: 294
 Score: 423.00 Matches: 91
 Percent Similarity: 52.3% Conservative: 58
 Best Local Similarity: 31.9% Mismatches: 118
 Query Match: 9.6% Indels: 18
 Gaps: 2
 DB: 3
 US-10-768-158-1 (1-2419) x Q2XV97_MACFA (1-294)
 QY 60 GAGTTCGAGACGAAGTACTTCGAGTTCCATGGCGTCCGCTGCGCCCTTCTGCGCGGG 119
 DB 6 AspTyrTyrGluAsnPheGluGluLeuHieGluValLeuMetTyrLysAspPheValLys 25
 QY 120 AAGATGGAGAGATCGCAACTTCGCGTCCGCGCCAGGACGTGGTGGATCGTCACTAC 179
 DB 26 TyrTrpAsnAspValGluThrPheGlnAlaArgProAspAspLeuValIleAlaThrTyr 45
 QY 180 CCAAGTCCGCGACCACTGTCGAGGAGGTGGTCTACTGTGTAGCCAGCGCGCTGAC 239
 DB 46 ProLysSerGlyThrTrpValSerGluIleAlaTyrMetIleTyrLysGluGlyAsp 65
 QY 240 CCGATGAGATCGGCTTGATGAACATCGACGACGCTCCGCTCTCGAGTACCCACAG 299
 DB 66 ValGluLysCysLysGluAspValIlePheAsnArgIleProPheLeuGluCysArgLys 85
 QY 300 CCG-----GGCTGGACATCATCAAGAACTGACCTCTCCCGCTCATCAAG 347
 DB 86 GluAspLeuMetAsnGlyValLysGlnLeuAspGluMetAsnSerProArgIleValLys 105
 QY 348 AGCAGCTGCGCTACCGCTTCTGCGCTCTGACCTCCACAAAGGAGACTCCAGGTATC 407
 DB 106 ThrHisLeuProGluLeuLeuProAlaSerPheTrpGluLysAsnCysLysIleIle 125
 QY 408 TATATGCTCGCAACCCCAAGGATGGTGGTCTTATTATTCAGTTCCACCGCTCTCG 467
 DB 126 TyrLeuCysArgAsnAlaLysAspValAlaValSerPheTyrTyrPhePheLeuMetVal 145
 QY 468 CGGACCATGAGCTACCGAGCACCTTCAAGAAATCTGCGCGAGGTTTATGATGATAAG 527
 DB 146 AlaGlyHisProAsnProGlySerPheProGluPheValGluLysPheMetGlnGlyGln 165
 QY 528 CTGGGCTACGGCTCTGGTTTGGACAGGTGACGAGTCTTGGGAGCACCGCATGCACTCG 587
 DB 166 ValProTyrGlySerTrpTyrLysHisValLysSerTrpTrpGluLysGluLysSerPro 185

```

QY 78 TTCGAGTTCATGGCGTGGCGTGGCGTCTCTCGCGGGAAGATGGAGAGATCGCC 137
Db 21 MetGluAlaAspGlyValProThrLeuLeuSerLysGluTrpTrpGluLysValCys 40
QY 138 AACTTCCCGTGGCGCCAGCAGCAGCAGTGGATCGTCACCTACCCCAAGTCGGGACACGAGC 197
Db 41 AenPheGlnAlaLysProAspAspLeuLeuAlaThrTyProLysSerGlyThrThr 60
QY 198 TTGCTGAGGAGTGGTCTACTGTGTGAGCCAGCGCGTGTGACCCCGATGAGATCGGCTTG 257
Db 61 TrpMetHisGluLeuLeuAspMetIleLeuAenAspGlyAspValGluLysCysIysArg 80
QY 258 ATGACATCGACGACGAGCTCCGCTCTGGAG-----TACCCA-----CAGCG 302
Db 81 AlaGlnThrLeuAspArgHisAlaPheLeuGluLeuLysPheProHisLysGluLysPro 100
QY 303 GGCCTGGACATCATCAAGGAATGACCTCTCCCGCCCTCATCAAGAGCAGCAGTCCCTTAC 362
Db 101 AspLeuGluPheValLeuGluMetSerSerProGlnLeuIleLysThrHisLeuProSer 120
QY 363 CGCTTTCTGCTCTGACTCCACAAATGAGAGACTCAAGGTCAAGTCTATATGCTGCGAAC 422
Db 121 HisLeuIleProProSerIleTrpLysGluAenCysLysIleValTyValAlaArgAen 140
QY 423 CCCAAGGATCTGGTGTCTATTATCAGTTCCACCGCTCTCTCGGACCATGAGCTAC 482
Db 141 ProlLysAspCysLeuValSerTyTrpHisPheHisArgMetAlaSerPheMetProAsp 160
QY 483 CGAGGACCTTTCAAGAAATTCGCGGAGGTTTATGAATGATTAAGCTGGCTACCGCTCC 542
Db 161 ProGlnAenLeuGluGluPheTyTrpGluLysPheMetSerGlyLysValValGlySer 180
QY 543 TGGTTGAGCAGTCGACGAGTTCCTGGAGACCGCATGAGCTCGAAGTGTCTTTTCTC 602
Db 181 TrpPheAspHisValLysGlyTrpTrpAlaAlaLysAspMetHisArgIleLeuTyLeu 200
QY 603 AGATGATGAGACATCATCGGACCTGTGTGAGATGGTGGACACTGGCCAGATTCCTG 662
Db 201 PheTyTrpGluAspIleLysLysAspProLysArgGluIleGluLysIleLeuLysPheLeu 220
QY 663 GGGGTGTCTGTGACAAGGCCAGCTGGAAGCCCTGACGAGCAGTCCACCATGCTGTG 722
Db 221 GluLysAspIleSerGluGluLeuLeuAenLysIleIleTyHisThrSerPheAspVal 240
QY 723 -----GACCAGTGTGCAACGCTGAGGCCCTCGCGCTGGGC----- 758
Db 241 MetLysGlnAenProMetThrAsnTyTrpThrLeuProThrSerIleMetAspHisSer 260
QY 759 -----CGGGAGAGATGGGCTGTGGAGGACATCTTACCGTCTCC 800
Db 261 IleSerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyPheThrValAla 280
QY 801 ATGAATGAGAAGTTTGACTGTGTTGATATAACAGAAGATGGGAAGTGTGACCTCACGTTT 860
Db 281 GlnAenGluGluPheAspLysAspTyTrpGlnLysLysMetAlaGlySerThrLeuThrPhe 300

```

RESULT 68

```

Q6XZC1_BRARE
ID Q6XZC1_BRARE PRELIMINARY; PRT; 304 AA.
AC Q6XZC1
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DE SUL1 sulfotransferase isoform 15.
DE SUL1 sulfotransferase isoform 4.
GN Name=sulf1st4; ORFNames=zgc:76922;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP NUCLEOTIDE SEQUENCE.

```

```

RA Liu M.-C., Liu C.-C., Sugahara T.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hong L.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hsieh F.,
RA Stapleton M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Brownstein M.J., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.H.,
RA Raba S.S., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A.C., Shevchenko V., Bouffard G.G.,
RA Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RX NIH MGC Project;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC
DR EMBL; AY196986; AAP5638.1; -; mRNA.
DR EMBL; BC066584; AAP6584.1; -; mRNA.
DR HSSP; P49891; 1AQU.
DR Ensemble; ENSDARG0000003181; Danio rerio.
DR ZFIN; ZDB-GENE-040426-2054; sulf1st4.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase 1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
DR Transferase.
SQ SEQUENCE 304 AA; 35556 MW; 20BE69D9DC409873 CRC64;

```

Alignment Scores:

Pred. No.:	5,52e-21	Length:	304
Score:	421.50	Matches:	98
Percent Similarity:	53.0%	Conservative:	51
Best Local Similarity:	34.9%	Mismatches:	111
Query Match:	9.6%	Indels:	21
DB:	2	Gaps:	6

US-10-768-158-1 (1-2419) x Q6XZC1_BRARE (1-304)

```

QY 78 TTCGAGTTCATGGCGTGGCGTGGCGTCTCTCGCGGGAAGATGGAGAGATCGCC 137
Db 21 PheAspPheGluGlyValPheLeuThrArgPhePheThrAspAsnTrpGluAenValLys 40
QY 138 AACTTCCCGTGGCGCCAGCAGCAGTGTGGATCGTCACCTACCCCAAGTCGGGACACGAGC 197
Db 41 AenPheGlnAlaArgProAspAspIleLeuIleAlaThrTyProLysAlaGlyThrThr 60
QY 198 TTGCTGAGGAGTGGTCTACTGTGTGAGCCAGCGCTGACCCCGATGAGATCGGCTTG 257
Db 61 TrpValSerTyTrpIleLeuAspLeuLeuTyTrpPheGlySerAspGluAenGlnThrSer--- 79
QY 258 ATGAACATCGACGACGAGCTCCGCTCGTCTGAG-----TACCCAAG 299
Db 80 GlnProIleValGlnArgValProPheLeuGluSerCysPheGlnGluPheSerThrIle 99

```

```
QY 300 CCGGCCTGGACATCATCAAGAACTG---ACCTCTCCCGCCTCATCAAGACCCACCTG 356
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
100 SerGlyThrGluMetAlaAspAenLeuProThrSerProArgLeuIleLysThrHisLeu 119
QY 357 CCCTACCGCTTCTCGCCCTCTGACCTCCACATGAGACTCAAGGTATCTATATGCT 416
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
120 ProValGlnLeuValProLysSerPheTrpGluGlnAenSerArgValValTyrValAla 139
QY 417 CCACACCCCAAGGATCTGGTGTCTTATTATCAGTTCCACCGCTCTCTGCGGACCATG 476
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
140 ArgAsnAlaLysAspAsnAlaValSerTyrPheHisPheAspArgMetAsnMetValGln 159
QY 477 AGCTACCGAGGACCTTCAAGAACTTCGCGGAGTTTATGAATGATAAGCTGGCTAC 536
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
160 ProAspProGlyAspTrpAspSerTyrLeuAspLysPheMetGlnGlyGlnAsnValPhe 179
QY 537 GGCTCTCTGTTGACAGCTGACAGCTCTGGAG---CACCGCATGGACTCGAACGTG 593
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
180 GlySerTrpPheAspHisValSerGlyTrpTrpGlnLysLysArgSerTyrProAsnMet 199
QY 594 CTTTTCCTCAAGTATGAACATGATCGGACCTGGTGACGATGGTGAGCAGCTGGCC 653
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
200 LeuTyrMetPheGluAspLeuSerGluAspThrGlyArgGluValAsnArgLeuGlyCys 219
QY 654 AGATTCTGGGGTCTCTGTGACAAAGCCAGCTGGAAGCCCTCACGAGCATGCCAC 713
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
220 SerPheLeuGlyLeuSerThrSerValGlnGluLysGluLysIleThrLysGlyValGln 239
QY 714 -----CAGCTGGTGACGAGCTGTCACAGCTGAGGCCCTGCC----- 752
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
240 PheAspAlaMetLysGlnAsnThrLeuIleAsnHisValThrIleProPheLeuAspCys 259
QY 753 -----GTGGCGCGGAAGAGTGTGGCTGTGAAGGACATCTTACCGTC 797
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
260 LysIleSerProPheMetArgLysGlyLysValGlyAspTrpLysSerHisPheThrVal 279
QY 798 TCCATGAATGAAAGTTTGTGCTGTATTAACAGAGATGGGAAAGTGTGACCTCAG 857
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
280 AlaGlnAenGluArgPheAspGluValTyrLysGlnLysMetLysAsnSerGlyValThr 299
QY 858 TTT 860
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
300 Phe 300

RESULT 69
ST1E1_RAT
ID ST1E1_RAT STANDARD; PRT; 295 AA.
AC P52844;
DT 01-OCT-1996, integrated into UniProtKB/Swiss-Prot.
DT 01-OCT-1996, sequence version 1.
DE Estrogen sulfotransferase, isoform 1 (EC 2.8.2.4) (EST-1)
DE (Sulfotransferase, estrogen-prefering) (Estrone sulfotransferase).
GN Name=Stultiel; Synonyms=Ste, Steel;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA].
RC STRAIN=Fischer 344; TISSUE=Liver;
RX MEDLINE=96305357; PubMed=8688469; DOI=10.1016/0167-4781(96)00065-6;
RA Rikke B.A., Roy A.K.
RT "Structural relationships among members of the mammalian
RT sulfotransferase gene family."
RL Biochim. Biophys. Acta 1307:331-338(1996).
CC -!- FUNCTION: Sulfation of estrone and estradiol. May control the
CC level of the estrogen receptor by sulfonylating free estradiol (By
CC similarity).
CC -!- CATALYTIC ACTIVITY: 3'-phosphoadenylyl sulfate + estrone =
CC adenosine 3',5'-bisphosphate + estrone 3-sulfate.
```

```
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasm (By similarity).
CC -!- SIMILARITY: Belongs to the sulfotransferase family.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
DR EMBL: U50204; AA07680.1; -; mRNA.
DR HSSP: P49891; 1B06.
DR SMR: P52844; 7-294.
DR Ensembl: ENSRNOG0000001957; Rattus norvegicus.
DR RGD: 3776; Ste.
DR InterPro: IPR000863; Sulfotransferase.
DR Pfam: PF00685; Sulfotransferase_1; 1.
DR ProDom: PD001218; Sulfotransferase; 1.
KW Lipid-binding; Steroid-binding; Transferase.
FT CHAIN 1 295
FT Estrogen sulfotransferase, isoform 1.
FT /FTID=PRO_0000085155.
FT NP_BIND 48 53
FT NP_BIND 130 138
FT NP_BIND 193 229
FT NP_BIND 257 259
FT ACT_SITE 108 108
FT Proton acceptor (By similarity).
SQ SEQUENCE 295 AA; 35509 MW; 696A12FDA923A12E CRC64;

Alignment Scores:
Pred. No.: 5.94e-21 Length: 295
Score: 421.00 Matches: 91
Percent Similarity: 53.0% Conservative: 60
Best Local Similarity: 31.9% Mismatches: 116
Query Match: 9.6% Indels: 18
DB: 1 Gaps: 3

US-10-768-158-1 (1-2419) x ST1E1_RAT (1-295)
QY 60 GAGTTCGAGACAGTACTTCGAGTTCCATCGCGTGGCGCTCTCTGCGCGCGG 119
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7 GluTyrTyrGluValPheGlyAspPheHisGlyPheLeuMetAspLysArgPheThrLys 26
QY 120 AAGATGAGGAGATCGCAACTTCGCGTGGCGGCCAGCGCTGTGTGATCGTCACTTAC 179
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
27 TyrTrpGluAspLleGluThrPheLeuAlaArgProAspAspLeuLeuIleValThrTyr 46
QY 180 CCCAAGTCCGCGCACGAGCTTCTGTCAGAGAGTGTGTACTTTGGTAGCCAGGCGCTAC 239
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
47 ProLysSerGlySerThrTrpIleSerGluIleValAspMetIleTyrLysGluGlyAsp 66
QY 240 CCGATGAGATCGGCTTGATGAACATCGACGAGCTCCCGGTCCTGGAGTACCCACAG 299
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
67 ValGluLysCysLysGluAspAlaLeuPheAsnArgIleProAspLeuGluCysArgAsn 86
QY 300 CCG-----GGCTGGACATCATCAAGGAAGTCACTCTCCCGGCTCATCAAG 347
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
87 GluAspLeuIleAsnGlyIleLysGlnLeuLysGluLysGluSerProArgIleValLys 106
QY 348 AGCCACCTGCGCTACCGCTTTCTGCTCCCTCTGACCTCCACAATGGAGACTCCAAGTCTATC 407
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
107 ThrHisLeuProAlaLysLeuLeuProAlaSerPheTrpGluLysAsnCysLysIleIle 126
QY 408 TATATGGCTCGCAACCCCAAGGATCTGGTGTGTCTTATTATTCAGTTCACCGCTCTCTG 467
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
127 TyrLeuCysArgAsnAlaLysAspValValValSerTyrTyrPhePheLeuIleMet 146
QY 468 CGGACCATGAGTACCAGGAGGACCTTTCAAGAAATTCGCCCGAGGTTTATGAATGATAAG 527
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
147 LysSerTyrProAsnProLysSerPheSerGluPheValGluLysPheMetGluGlyGln 166
QY 528 CTGGGCTACGGCTCTCTGGTTTGGACACGTGCAGAGGTTCTGGGACACCGCATGGACTCG 587
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
167 ValProTyrGlySerTyrTrpAspHisValLysSerTrpTrpGluLysSerLysAsnSer 186
QY 588 AACGTGCTTTTCTCAAGTATGAAGACATGCATCGGACCTGGTGACGATCGTGGAGCAG 647
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```


Db 187 ArgValLeuPheMetPheTyrGluAspMetLysGluAspIleArgArgGluValLys 206
 QY 648 CTGGCCACATTCCTGGGGGTGCTGTGACAAAGCCAGCTGGAAGCCCTGACGAGCAC 707
 Db 207 LeuIleGluPheLeuGluArgAspProSerAlaGluLeuValAspArgIleIleGlnHis 226
 QY 708 TGC-----CACCAGCTGTGACACAGCTGCTGCAACGCTGAGGCCCTGCC----- 752
 Db 227 ThrSerPheGlnGluMetLysAsnAsnProCysThrAsnTyrSerMetLeuProGluThr 246
 QY 753 -----GTGGCCCGGGAAGAGTGGCTGTGGAAGAC 785
 Db 247 MetIleAspLeuLysValSerProPheMetArgLysGlyIleValGlyAspTrpArgAsn 266
 QY 786 ATCTTCACCGCTCCATGAATGAGAGTTGACITGGTGTATTAACACAAAGATGGAAAG 845
 Db 267 HisPheProGluAlaLeuArgGluArgPheGluGluHisTyrGlnArgHisMetLysAsp 286
 QY 846 TGTGACCTCACGTTT 860
 Db 287 CysProValLysPhe 291

RESULT 70

Q95JCS_ORNAN
 ID Q95JCS_ORNAN PRELIMINARY; PRT; 295 AA.
 AC Q95JCS57
 DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
 DT 01-DEC-2001, sequence version 1.
 DT 07-FEB-2006, entry version 15.
 DE Sulfotransferase SULF1A.
 OS Ornithorhynchus anatinus (Duckbill platypus).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Monotremata; Ornithorhynchidae; Ornithorhynchus.
 OX NCBI_TaxID=9258;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Liver;
 RA Bolton-Grob R.M., McManus M.E.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License

CC -----
 DR EMBL; AV044182; AAK72405.1; -; mRNA.
 DR HSSP; P50224; 1CJM.
 DR SMR; Q95JCS5; 8-235.
 DR GO; GO:0008146; F:sulfotransferase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR InterPro; IPR000863; Sulfotransferase.
 DR Pfam; PF00685; Sulfotransferase 1; 1.
 DR ProDom; PD001218; Sulfotransferase; 1.
 KW Transferase.
 SQ SEQUENCE 295 AA; 34189 MW; 8528DA3ECC8833B7 CRC64;

Alignment Scores:

Pred. No.: 7e-21 Length: 295
 Score: 420.00 Matches: 95
 Percent Similarity: 54.2% Conservatives: 61
 Best Local Similarity: 33.0% Mismatches: 113
 Query Match: 9.5% Indels: 19
 DB: 2 Gaps: 3

US-10-768-158-1 (1-2419) x Q95JCS_ORNAN (1-295)

QY 53 CCCGGGGAGTTTCGAGAGCAAGTACTTCGAGTTCCATGGCGTGGCGCTTCTG 112
 Db 5 ProAspThrSerArgProProValArgVal-AsnGlyIleProLeuIleLysTyrPh 24
 QY 113 CCGCGGGAAGATGGAGGAGATCGCAACTTCCCGTGGCGCCAGCAGCTGTGATCGT 172
 Db 24 eAlaGluAsnLeuGlyAlaLeuGluSerLeuArgValGlnProSerAspLeuValSe 44
 QY 173 CACCTACCCCAAGTCCGGCACCAGCTTGTGTCAGGAGGTGGTCTACTTGGTGAGCCAGGG 232

Db 44 rThrTyrProLysSerGlyThrTrpValSerGluIleLeuAspMetIleTyrGlnG 64
 QY 233 CGCTGACCCCATGATCGCTTGTATGACATCGACGAGCAGCTCCGGTCTCGGAGTA 292
 Db 64 yGlyAspLeuGluLysCysGlnArgAlaProValPheLeuArgValProPheLeuGluPh 84
 QY 293 CCCACAGCG-----GGCCTGGACATCATCAAGGAAGCTGACCTCTCCCGCCT 340
 Db 84 eSerIleProGlyMetProSerGlyMetGluThrLeuLysAspThrProSerProA 104
 QY 341 CATCAAGAGCCACTGCTCCCTTCTGCTCCCTCTGACCTCCCAATGGAGACTCAA 400
 Db 104 uLeuLysThrHisLeuProLeuAlaLeuLeuProLysValLeuLeuAspGlnLysVal 124
 QY 401 GGTCACTATATGCTCGCAACCCCAAGGATCTGGTGTCTTATATATCATGTTCCACCG 460
 Db 124 sValIleTyrIleAlaArgAsnAlaLysAspValAlaValSerTyrTyrHisPheTyrAr 144
 QY 461 CTCTCTCGGACCATGAGCTACCGAGCACCTTCAAGAATTCTGCCGAGGTTTATGAA 520
 Db 144 gMetAlaLysValHisProAspProGlyThrTrpGluThrPheLeuGluAlaPheLysAl 164
 QY 521 TGATAAGCTGGCTACCGCTCTCTGTTGACACGTCGAGAGTTCCTGGAGCACCGCAT 580
 Db 164 aGlyGlnValCysTyrGlySerTrpTyrGlnHisValGlnAspTrpTrpGluLeuArgLy 184
 QY 581 GGACTCGNACGTGCTTTCTCAAGATGATGACATGCATCGGAGCTGGTGAGCATGGT 640
 Db 184 sGlnGlnProValLeuTyrLeuPheGluAspIleLysGluAspProLysArgGluI 204
 QY 641 GGACGAGCTGCCAGATTCTCTGGGGGTCTCTGTGACAAGGCCAGCTGGAAGCCCTG-- 698
 Db 204 eArgLysIleMetGluPheIleAspGlnProValSerGluGluLeuLeuAspArgL 224
 QY 699 -----ACGAGCAGCTGCCACCGCTGGTGGACAGCTGTGCTGCAACGCTGAGGCCCTGCC 751
 Db 224 lGlnGluThrSerPheLysLysMetLysGlnAsnProMetThrAsnTyrSerSerValPr 244
 QY 752 C-----GTGGCCGGGGAAGAGTTGGGCTGTG 778
 Db 244 oSerHisIleMetAspHisGluValSerProPheMetArgLysGlyThrProGlyAspTr 264
 QY 779 GAAGGACATCTTACCGCTCTCCATGAATGAGAAGTTTGACTTGTGTATATAACAAGAAT 838
 Db 264 pLysAsnHisPheThrValAlaGlnAsnGluLeuPheAspValAspTyrAlaGluLysMe 284
 QY 839 GGGAAAGTGTGACCTCACGTTT 860
 Db 284 tAlaGlySerGluLeuHisPhe 291

RESULT 71

Q3HYKO_STRPU
 ID Q3HYKO_STRPU PRELIMINARY; PRT; 285 AA.
 AC Q3HYKO;
 DT 08-NOV-2005, integrated into UniProtKB/TrEMBL.
 DT 08-NOV-2005, sequence version 1.
 DT 07-FEB-2006, entry version 3.
 DE Sulfotransferase.
 OS Strongylocentrotus purpuratus (Purple sea urchin).
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
 OC Echinoida; Euechinoida; Echinacea; Echinoida; Strongylocentrotidae;
 OC Strongylocentrotus.
 OX NCBI_TaxID=7668;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Heyland A., Moroz L.L., Price D.A., Bodnarova M.;
 RT "Thyroid hormone metabolism and thyroid peroxidase function in non-
 chordate animals.";
 RL Submitted (AUG-2005) to the EMBL/GenBank/DBJ databases.
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License


```

CC      EMBL; DQ176319; ABA41638.1; -, mRNA.
DR      GO; GO:0008146; F:sulfotransferase activity; IEA.
DR      GO; GO:0016740; F:transferase activity; IEA.
DR      InterPro; IPR000863; Sulfotransferase.
DR      Pfam; PF00685; Sulfotransferase_1; 1.
DR      ProDom; PD001218; Sulfotransferase; 1.
KW      Transferase.
SQ      SEQUENCE 285 AA; 33378 MW; B27DCDDBI7F06518 CRC64;

Alignment Scores:
Pred. No.:      7,52e-21      Length:      285
Score:          419.50      Matches:      94
Percent Similarity: 53.7%      Conservative: 59
Best Local Similarity: 33.0%      Mismatches: 105
Query Match:      9.5%      Indels:      27
DB:              2      Gaps:        5

US-10-768-158-1 (1-2419) x Q3HYKO_STRPU (1-285)
Qy      99 CTGCGCCCTTCTGCGCGGGAAG---ATGGAGGAGATGCCCAACTTCCGGTGGCGCC 155
      :::||| ||| ::: ::: ::: ::: ::: ::: ::: ::: ::: :::
Db      1 MetProlyserPheThrProLysArgPheLeuAspAspLeuLysAsnPheGluValLysGly 20
Qy      156 AGCAGCTGTGGATCGTCACTACCCCAAGTCCGCCAGCAGCTGCTGTCGAGAGTGTC 215
      ||| ::::||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      21 AspAspThrTyrLeuIleThrTrpProLysSerGlyThrTrpMetGlnAsnIleLeu 40
Qy      216 TACTTGGTGAGCCGGGCGCTGACCCCGATGAGATCGGCTTGATGACATCAGCAGCAG 275
      ||| ::::||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      41 ThrLeuIlePheAlaLysGlyAspMetAspAlaValArgGluLysHisLeuPheLysArg 60
Qy      276 CTCGCGGCTCTGGAGTACCCACAG-----CGGGC 305
      ||| ::::||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      61 ValProPheLeuGluMetProLysGlyPheAspTyrLysLysAlaGluAspThrGly 80
Qy      306 CTG---GACATCATCAAGAACTGACCTCTCCCGCCTCATCAAGAGCACCTGCCCTAC 362
      ||| ::::||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      81 LeuTyrGluIleValArgAsnValProSerProArgLeuLeuLysThrGlnLeuProPro 100
Qy      363 CGCTTTCTGCGCTCTGACCTCCACAACTGGAGACTCCAGGTCATCATATGCTCGCAC 422
      ||| ::::||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      101 ProPheLeuProThrGlnIleHisGluLysLysProLysIleValTyrValAlaArgAsn 120
Qy      423 CCCAAGGATCTGGTGGTCTCTATTATCAGTTTCCACCGCTCTCTCGGACCACTAGCTAC 482
      ||| ::::||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      121 ProLysAspAlaAlaValSerTyrPheHisPheCysAsnValSerProAsnLeuProGln 140
Qy      483 CGAGGCACCTTTCAAGAAATCTGCGCGAGGTTTATGAATGATGAAGCTGGCTACGGCTCC 542
      ::: ::::||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      141 TyrArgAspTyrAsnAspPhePheIleAspPheCysAsnAspSerIleProArgGlySer 160
Qy      543 TGGTTTGAGCAGCTGCAGGATCTGGGAGCACCCTGAGCTCGAAAGTGCTTTTCTC 602
      ||| ::::||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      161 TrpPheGluAsnValLeuTyrTrpTrpAsnLysArgHisGluSerAsnValLeuPheIle 180
Qy      603 AAGTATGAAGACATGCATCGGACCTGTGTGACGATGGTGGACAGCTGCCAGATTCTCTG 662
      ||| ::::||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      181 ThrTyrGluGluMetLysGlnAspLeuArgGlySerValValArgValCysAspPheLeu 200
Qy      663 GGGGTGTCTCTGTGACAAAGCCAGCTGGAAGCCCTGACGGAGCACTGC-----CAC 713
      ||| ::::||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      201 GlyLysGluLeuSerAspAspIleIleAspValIleThrGluAsnSerThrPheAsnAla 220
Qy      714 CAGCTGTGGACCGAGTGTGTGCNACCGCTGAGGCCCTGCCCGT-----755
      ||| ::::||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      221 MetLysLysAspProThrAlaAsnProAspSerLeuLeuValPheLysGluAlaLys 240
Qy      756 -----GGCGGGGAAGAGTTGGCTGTGGAGGAGTGGAAAGTGTGACCTCAGC 857
      ::: ::::||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      241 GlnLysArgSerPheLeuArgLysGlyGluValGlyAspTrpLysAsnHisPheThrVal 260
Qy      798 TCCATGAATGAGAAGTTTGACTTGGTGTATAAACAGAAGATGGGAAAGTGTGACCTCAGC 857

```


QY 697 CTGGAGCCCTG-----ACGAGACAGTCCACAGCTGGTGGACCAAGTGGTCAAC 737
 Db 214 ValGluIysIleAlaSerGlnThrSerPheIysAlaMetIysGlnAsnGluLeuSerAsn 233
 QY 738 GCTGAGGCCCTGGCC-----GTGGCCCGGGGA 764
 Db 234 TyrSerMetValProSerSerValMetAspHisSerIleSerProPheMetArgIysGly 253
 QY 765 AGAGTTGGGTGTGAAGGACATCTTCAACGCTCCATGAATGAGAGATTGACGTGTG 824
 Db 254 ValCysGlyAspTrpIysAsnGlnPheThrValAlaGlnAsnGluIysPheAspGluTyr 273
 QY 825 TATAACAGAGATGGGAAGTGTGACCTCAGCTTT 860
 Db 274 TyrGlnArgGluMetSerAspGlyAlaLeuSerPhe 285

RESULT 75
 Q88EV4_XENLA
 ID Q88EV4_XENLA PRELIMINARY; PRT; 276 AA.
 AC Q68EV4;
 DT 11-OCT-2004, integrated into UniProtKB/TrEMBL.
 DT 11-OCT-2004, sequence version 1.
 DT 07-FEB-2006, entry version 12.
 DE MGC84291 protein.
 GN Name=MGC84291;
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae;
 OC Xenopodinae; Xenopus; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Brain;
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.,
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 RT initiative."
 RL Dev. Dyn. 225:384-391 (2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Brain;
 RA Klein S., Gerhard D.S.;
 RL Submitted (Aug-2004) to the EMBL/GenBank/DBJ databases.
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs License

DR EMBL; BC080096.1; -, mRNA.
 DR GO; GO:0008146; F:sulfotransferase activity; IEA.
 DR InterPro; IPR000863; Sulfotransferase.
 DR Pfam; PF00685; Sulfotransfer_1; 1.
 DR ProDom; PD001218; Sulfotransferase; 1.
 SQ SEQUENCE 276 AA; 31979 MW; 4A2B98654EF33783 CRC64;
 Alignment Scores: 1.69e-20 Length: 276
 Pred. No.: 414.50 Matches: 96
 Score: 51.2% Conservative: 48
 Percent Similarity: 51.2% Mismatches: 100
 Best Local Similarity: 34.2% Indels: 37
 Query Match: 9.4% Gaps: 6
 DB: 2
 US-10-768-158-1 (1-2419) x Q68EV4_XENLA (1-276)
 QY 99 CTGCGCCCTTCTGCGCGGGAAGATGGAGAGATGCCAACTTCCCGTGGCGCCAGC 158
 Db 4 LeuGlyProPheAlaGlu--AsnTrpGluAsnValGluIysPheGlnAlaArgProAsp 22
 QY 159 GACGTGTGGATCGTCACCTACCCCAAGTCGCGCACCGCTGCTGCGAGAGGTGGTCTAC 218
 Db 23 AspLeuLeuIleCysThrTyrProIysSerGlyThrThrTrpIleCysGluIleVal--- 41
 QY 219 TTGTGTGACCCAGGCGGTGACCCCGATCGATCGCTGATGAAC----- 263
 Db 42 -----AspGlnIleLeuAlaValAsnAsnAlaGluGlyCys 53
 QY 264 -----ATCGACGACGAGCTCCCGCTCTGGAGTACCCACAGCCG----- 302
 Db 54 LysAsnAlaAlaIlePheGluArgValProPheLeuGluTyrAlaValProAsnMetIle 73
 QY 303 ---GGCTGGACATCATCAAGAACTGACCTCTCCCGCTCATCAAGAGCACCTGCGCC 359
 Db 74 SerGlyThrAlaAlaLeuAspGlnArgAlaSerProArgIleIleLysThrHisLeuPro 93
 QY 360 TACCGCTTTCGCTCTGACCTCCACATCGAGCTCCAGGTCCAGGTCACTATATGGCTGCG 419
 Db 94 ValGluLeuLeuProLysSerPheTrpAspAsnLysValLysIleIleTyrValAlaArg 113
 QY 420 AACCCCAAGGATCTGGTGTCTATTATCAGTTCCACCGCTCTCTCGGACCATGAGC 479
 Db 114 AsnAlaLysAspValAlaValSerTyrTyrHisPheTyrGlnMetAlaIleValHisPro 133
 QY 480 TACCGAGCACCTTCAAGAAATTCGCGGAGGTTTATGAATGATTAAGCTGGGTACCGCC 539
 Db 134 GluProGlyThrTrpAspPheLeuAspSerTyrIleGluGlyLysValCysPheGly 153
 QY 540 TCCTGGTTTGGACGACGTCAGAGTTCTGGGAGCACCGCATGGAGCTCGAAGCTGCTTTT 599
 Db 154 ProTrpSerThrHisValLysValGlyTyrTrpGlnMetAlaLysLysTrpAspValLeuTyr 173
 QY 600 CTCAGATGATGAACATGCATCGGACCTGCTGACGATGGTGGAGCAGCTGGCCAGATTC 659
 Db 174 LeuPheTyrGluAspMetLeuGluAspLeuThrArgGluIleArgLysValValLysPhe 193
 QY 660 CTGGGGGTGTCCTGTGACAAAGCCCGACGTGAA-----GCCCTGACGAGGACCTGCG 710
 Db 194 MetGlyLysAspLeuSerLysGluLeuValGluLysIleAlaSerLeuThrSerPheLys 213
 QY 711 CACGACGTGGTGACGACGTGCGAAGCTGAGCCCTGAGCCCTGCGCC----- 752
 Db 214 AlaMetLysGluAsnLysAsnSerAsnTyrThrThrValProSerSerValMetAspHis 233
 QY 753 -----GTGGCGCGGGGAAGAGTTGGGTGTGGAGAGACATCTTACCGTC 797
 Db 234 SerIleSerProPheMetArgLysGlyValCysGlyAspTrpLysAsnGlnPheSerVal 253
 QY 798 TCCATGAATGAGAGTTTTCAGCTTGGTGATATAACAGAGATGGGAAAGTGTGACCTCAGC 857
 Db 254 AlaGlnAsnGluLysPheAspLysTyrTyrGlnArgGluMetAlaAspCysAspLeuSer 273

```
QY 858 TTT 860
Db 274 Phe 274

RESULT 76
O35401_MOUSE PRELIMINARY; PRT; 295 AA.
AC O35401_MOUSE
DT 01-JAN-1998, integrated into UniProtKB/TrEMBL.
DT 01-JAN-1998, sequence version 1.
DT 07-FEB-2006, entry version 28.
DE Amine N-sulfotransferase.
GN Name=Sultid1; Synonyms=SULT-N;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_TaxID=10090;
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98321187; PubMed=9647753; DOI=10.1006/bbr.1998.8872;
RA Sakakibara Y., Yanagisawa K., Takami Y., Nakayama T., Suiko M.,
RA Liu M.-C.;
RT "Molecular cloning, expression, and functional characterization of
RT novel mouse sulfotransferases.";
RL Biochem. Biophys. Res. Commun. 247:681-686(1998).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF026073; AAC69919.1; -; mRNA.
DR PIR; JEO197; JEO197.
DR HSSP; P49888; 1HY3.
DR Ensembl; ENSMUSG0000029273; Mus musculus.
DR MGI; MGI:1926341; Sultid1.
DR GO; GO:0004062; F-aryl sulfotransferase activity; IDA.
DR GO; GO:0000103; F-sulfate assimilation; IDA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransfer_1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Transferase.
SQ SEQUENCE 295 AA; 35099 MW; 17FE3E02AD6E1269 CRC64;

Alignment Scores:
Pred. No.: 2.2e-20 Length: 295
Score: 413.00 Matches: 83
Percent Similarity: 55.7% Conservative: 63
Best Local Similarity: 31.7% Mismatches: 98
Query Match: 9.4% Indels: 18
DB: 2 Gaps: 2

US-10-768-158-1 (1-2419) x O35401_MOUSE (1-295)
QY 129 GAGATCGCAACTTCCGGTGGCGCCGACGAGTGTGATCGTCACCTACCCCAAGTCC 188
Db 30 GlnValGluSerPheGluAlaArgProAspIleLeuIleSerThrTyrrProlySer 49
QY 189 GGCACGAGTGTCTGACGAGGTGTCTACTTGTGTGAGCCAGGGCGCTGACCCCGATGAG 248
Db 50 GlyThrThrTrpValSerGluIleLeuAspIleIleTyrrAsnAsnGlyAspAlaGluLys 69
QY 249 ATCGGCTTTGTAACATCGACGAGCAGCTCCGGTCTCGGAGTATCCCAACGCGGCGCTG 308
Db 70 CysLysArgAspAlaIleTyrrLysArgValProPheMetGluLeuIleIleProGlyIle 89
QY 309 -----GACATCATCAGAACTGACCTCTCCCGCTCATCAAGCCACCTG 356
Db 90 ThrAsnGlyValGluMetLeuAsnMetProSerProArgIleValIleThrHisLeu 109
QY 357 CCTACCGCTTCTCCCTCTGACCTCCACATGGAGACTCCAGGTCTATATATGCT 416
Db 110 ProValGlnLeuLeuProSerSerPheThrLysAsnAspCysLysIleIleIleTyrrValAla 129
```

```
QY 417 CGAACCCCAAGGATCTGGTGTCTTATATATCAGTTCCACCGCTCTCTCGGACCATG 476
Db 130 ArgAsnAlaLysAspValValSerTyrrTyrrPheTyrrGlnMetAlaLysIleHis 149
QY 477 AGTACCGAGGCACCTTTCAAGAAATTCGCGGAGGTTTATGAATGATAAGCTGGGTAC 536
Db 150 ProGluProGlyThrTrpGluPheLeuGluLysPheMetAlaGlyGlnValSerPhe 169
QY 537 GGCTCTGTTGTTGAGCAGCTGACAGAGTTCTGGGAGCACCGCATGGACTCGAAGCTGTT 596
Db 170 GlyProTrpTyrrAspHisValLysSerTrpTrpGluLysArgLysGluTyrrArgIleLeu 189
QY 597 TTTCTCAAGTATGAACATCATCGGACCTGTCAGCATGGTGGAGCAGCTGCCACAGA 656
Db 190 TyrLeuPheTyrrGluAspMetLysGluAsnProLysCysGluIleGlnGlnIleLeuLys 209
QY 657 TTCCTGGGGGTCTCTGTGACAAAGCCAGCTGGAAGCCCTGACGAGCAGCTGCCAC--- 713
Db 210 TyrLeuGluLysAspIleProGluGluLeuAsnLysIleLeuTyrrHisSerSerPhe 229
QY 714 -----CAGCTGGTGGACCAAGTGTGCTGC 734
Db 230 SerValMetLysGluAsnProSerAlaAsnTyrrThrMetMetLysGluGluMetAsp 249
QY 735 AACCTGAGGCCCTGCCCGTGGCGGAGAGAGTTGGCTGTGGAGGACATCTTCACC 794
Db 250 HisSerValSerProPheMetArgLysGlyIleSerGlyAspTrpLysAsnGlnPheThr 269
QY 795 GTCTCCCATGAATGAGAAGTTTGACTGTGTATAAACAGAGAAGATGGGAAAGTGTGACCTC 854
Db 270 ValAlaGlnTyrrGluLysPheGluGluAspTyrrValLysLysMetGluAspSerThrLeu 289
QY 855 ACCTTT 860
Db 290 LysPhe 291

RESULT 77
QSEAWO_XENLA PRELIMINARY; PRT; 303 AA.
AC QSEAWO;
DT 15-MAR-2005, integrated into UniProtKB/TrEMBL.
DT 15-MAR-2005, sequence version 1.
DT 07-FEB-2006, entry version 8.
DE Hypothetical protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Egg;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Egg;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
```

RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.W., Marra W.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 [3]
 RN NUCLEOTIDE SEQUENCE.
 RP TISSUE-Specific.
 RC Klein S., Gerhard D.S.;
 RA Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
 RL
 CC
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License

 DR EMBL; BC090224; AAH90224.1; -; mRNA.
 DR GO; GO:0008146; F:sulfolipase activity; IEA.
 DR InterPro; IPR000863; Sulfolipase.
 DR Pfam; PF00685; Sulfolipase 1; 1.
 DR ProDom; PD001218; Sulfolipase; 1.
 DR Hypothetical protein.
 SQ SEQUENCE 303 AA; 35647 MW; B940559678751DA1 CRC64;

 Alignment Scores:
 Pred. No.: 2,21e-20 Length: 303
 Score: 413.00 Matches: 89
 Percent Similarity: 52.0% Conservative: 55
 Best Local Similarity: 32.1% Mismatches: 115
 Query Match: 9.4% Indels: 18
 DB: 2 Gaps: 3

 US-10-768-158-1 (1-2419) x Q5EAWO_XENLA (1-303)
 Qy 84 TTCCATGCGTGGCGCTGCGCCCTCTGCGCGGGAGATGGAGGATCGCAACTTC 143
 Db 23 PheargGlyValProMetAlaLysMetAsnValGluAsnTrpGluProValGluTrpPhe 42
 Qy 144 CCGGTGGCGCCAGCGAGCTGTGGATCGTCACCTACCCCAAGTCCGGGACGAGTTGTG 203
 Db 43 GlnAlaArgHisAspAspValValLeuAlaThrTyProLysAlaGlyThrTrpVal 62
 Qy 204 CAGGAGGTGGTCTACTTGGTGAGCCAGGCGCTGACCCCGATGAGATCGCTTGATGAAC 263
 Db 63 SerGluIleLeuAspMetIleTyraenGlyValAspLeuGluLysCysGlnArgAspThr 82
 Qy 264 ATCGACGAGCAGCTCCCGGTCTCGGAGTACCCACACGCGGGCGTG-----GAC 311
 Db 83 IleTyraenArgValProTyMetGluIleArgIleProGlyMetProSerGlyValAsp 102
 Qy 312 ATCATCAAGGAACCTGACCTTCCCGGCTCATCAAGAGCCACCTGCGCTACCGTTCTGT 371
 Db 103 GlnLeuGluLeuLeuAlaSerProArgLeuIleLysThrHisLeuProIleGlnLeuMet 122
 Qy 372 CCCTCTGACCTCCCAATGGAGACTCCAGGTCATCTATATGGCTCGCAACCCAGGAT 431
 Db 123 ProGluSerPheTrpGluLysLysCysLysValIleTyraenValAlaArgAsnAlaLysAsp 142
 Qy 432 CTGGTGGTGTCTTATTATTCAGTTCCACCGCTCTCTCGCGGACCATGAGTACGAGGCACC 491
 Db 143 ValAlaValSerTyPhePheHisGlnMetValLysAlaLeuProAspProGlyPro 162
 Qy 492 TTCAAGGAATCTCGCGGAGGTTTATGAATAGTAAAGCTGGGTACCGCTCTCGTTTGTAG 551
 Db 163 TrpAspLysPheLeuAlaAspTyMetAsnGlyThrValSerTyGlySerTrpTyAsp 182
 Qy 552 CACGTGACGAGGTTCTCGGAGCACCGCATGGACTCGAACGTGCTTTTCTCAAGTATGAA 611
 Db 183 HisValLysGlyTrpTrpGluLysArgLysGlnTyraenTyraenLeuPheTyraenGlu 202
 Qy 612 GACATGCATCGGAGCTGGTGTACGATGGTGGAGCAGCTGGCAGATTCCTGGGGGTGTC 671

DR InterPro: IPR000863; Sulfotransferase.
DR Pfam: PF00685; Sulfotransferase_1; 1.
DR ProDom: PD001218; Sulfotransferase; 1.
KW Transferase.
SQ SEQUENCE 304 AA; 35797 MW; 9CF4A4B63713B977 CRC64;
Alignment Scores:
Pred. No.: 2.4e-20 Length: 304
Score: 412.50 Matches: 89
Percent Similarity: 52.0% Conservative: 56
Best Local Similarity: 31.9% Mismatches: 115
Query Match: 9.4% Indels: 19
DB: 2 Gaps: 3
US-10-768-158-1 (1-2419) x Q80VR3_MOUSE (1-304)
QY 81 GAGTTCATGGCTGCGGCTGCGGCTTCTCGCGCGGAAGATGGAGGAGATCGCCAAC 140
DB 22 GluValAsnGlyIleLeuMetSerLysMetMetSerGluAsnTrpAspLysIleTrpAsn 41
QY 141 TTCGCGGTGGCGCCAGCGACGTGTGGATGTCACCTACCCCAATCCGCGCACACGCTTG 200
DB 42 PheGlnAlaLysProAspAspLeuIleAlaThrTyrAlaLysAlaGlyThrThrTrp 61
QY 201 CTGACGAGGTGCTTACTTGTGTGACCGCGGCTGACCCCGATGAGATCGGCTTGATG 260
DB 62 ThrGlnGluIleValAspMetIleGlnAsnAspGlyAspValGlnLysCysGlnArgAla 81
QY 261 AACATCGACGACGACCTCCGCTCTCGGAGTACCCACACGCG-----GGC 305
DB 82 AsnThrTyrAspArgHisProPheIleGluThrLeuProProLeuAsnSerGly 101
QY 306 CTGGACATCATCAAGAACTACCTCCCGGCTCATCAAGAGCCACTGCGCTACCGC 365
DB 102 LeuAspLeuAlaAsnLysMetProSerProArgThrLeuLysThrHisLeuProValGln 121
QY 366 TTTCGCGCTCTGACCTCCACATGAGACTCCAGGTCATCTATATGCTCGCGAACCC 425
DB 122 MetLeuProProSerPheTrpLysGluAsnSerLysIleIleTyrValAlaArgAsnAla 141
QY 426 AAGGATCTGGTGTCTATTATCAGTTCCACCGCTCTCTCGCGACCATGAGCTACCGA 485
DB 142 LysAspCysLeuValSerTyrTrpPheSerArgMetAsnLysMetLeuProAspPro 161
QY 486 GGCACCTTCAAGAACTTCGCGGAGGTTTATGAATGATAAGCTGGGCTACGCTCTCGG 545
DB 162 GlyThrLeuGlyGluTyrIleGluThrPheLysAlaGlyLysValLeuTrpGlySerTrp 181
QY 546 TTGAGCAGCTGCAGGAGTTCTGGGAGCACCGCATGGACTGCAAGCTGCTTTTCTCAAG 605
DB 182 TyrAspHisValLysGlyTrpTrpAspValLysAspLysHisArgIleLeuTyrLeuPhe 201
QY 606 TATGAACATGTCATCGGACCTGGTGACGATGGTGGAGAGCTGGCGAGATCTCTCGGG 665
DB 202 TyrGluAspMetLysGluAspProLysArgGluIleLysLysIleValLysPheLeuGlu 221
QY 666 GTGCTCTGTCAAGCCCGACGCTGGAAGCCCTGACGAGACCTGCCACCGACTGGTG--- 722
DB 222 LysAspIleSerGluGluValLeuAsnLysIleIleHisThrSerPheAspValMet 241
QY 723 -----GACAGTGTGCAAGCTGAGGCCCTGCGCC----- 752
DB 242 LysGlnAsnProMetAlaAsnTyrThrThrLeuProSerSerIleMetAspHisSerIle 261
QY 753 -----GTGGGCGGGGAGAGTGTGGCTGTGGAGGAGCATCTTCACCGCTCCCATG 803
DB 262 SerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAlaGln 281
QY 804 AATGAGAAGTTTGACTTGGTGTTATTAACAGAGATGGGAAGTGTGACTCACGTTT 860
DB 282 SerGluAspPheAspGluAspTyrArgLysLysMetAlaGlySerThrIleThrPhe 300

RESULT 79

STIAL BOVIN
ID STIAL_BOVIN STANDARD; PRT; 294 AA.
AC P50227;
DT 01-OCT-1996, integrated into UniProtKB/Swiss-Prot.
DT 01-OCT-1996, sequence version 1.
DT 07-FEB-2006, entry version 43.
DE Sulfotransferase 1A1 (EC 2.8.2.1) (Aryl sulfotransferase) (Phenol
DE sulfotransferase) (Phenol-sulfating phenol sulfotransferase) (P-PST).
GN Name=SULT1A1; Synonyms=STP;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE, AND PARTIAL PROTEIN SEQUENCE.
RC TISSUE=Tracheobronchial;
RX MEDLINE=96003918; PubMed=7575456;
RA Schauss S.J., Henry T., Palmatier R., Halvorson L., Dannenbring R.,
RA Beckmann J.D.;
RT "Characterization of bovine tracheobronchial phenol sulphotransferase
RT cDNA and detection of mRNA regulation by cortisol.";
RL Biochem. J. 311:209-217(1995).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=8890738; DOI=10.1016/0378-1119(96)00083-2;
RA Nonnenan D.J., Shibuya H., Johnson G.S.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Catalyzes the sulfate conjugation of catecholamines and
CC of phenolic drugs.
CC -!- CATALYTIC ACTIVITY: 3'-phosphoadenylyl sulfate + a phenol =
CC adenosine 3',5'-bisphosphate + an aryl sulfate.
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasm.
CC -!- TISSUE SPECIFICITY: Distal lung parenchyma.
CC -!- SIMILARITY: Belongs to the sulfotransferase family.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
EMBL: U35253; AA85510.1; -; mRNA.
EMBL: U34753; AAC48677.1; -; Genomic DNA.
EMBL: L33828; AA56789.1; -; Genomic DNA.
FIR: JC5000; JC5000.
HSSP: P50224; 1CUM.
DR SMR; P50227; 8-284.
DR InterPro: IPR000863; Sulfotransferase.
DR Pfam: PF00685; Sulfotransferase_1; 1.
DR ProDom: PD001218; Sulfotransferase; 1.
KW Catecholamine metabolism; Direct protein sequencing; Lipid metabolism;
KW Steroid metabolism; Transferase.
FT CHAIN 1 294 Sulfotransferase 1A1.
FT /FTID=PRO_0000085125.
FT NP_BIND 48 53 PAPS (By similarity).
FT NP_BIND 130 138 PAPS (By similarity).
FT NP_BIND 193 229 PAPS (By similarity).
FT NP_BIND 257 259 PAPS (By similarity).
FT ACT_SITE 108 108 Proton acceptor (By similarity).
SQ SEQUENCE 294 AA; 34017 MW; 8ADE567D47E69737 CRC64;
Alignment Scores:
Pred. No.: 2.58e-20 Length: 294
Score: 412.00 Matches: 91
Percent Similarity: 53.5% Conservative: 56
Best Local Similarity: 33.1% Mismatches: 108
Query Match: 9.4% Indels: 20
DB: 1 Gaps: 4

Db 63 SerAspGlyAspThrAspLysSerLysArgAspAlaIleHisMetLysValProMetLeu 82
 QY 288 GAGTACCCACAGCGGGCTGGAC-----ATCATCAAGAACTGACCTCT 332
 Db 83 GluPheSerAlaProGlyGlnValAlaSerGlySerLeuValLeuGluSerValProSer 102
 QY 333 CCCCGCTCATCAAGACCACCTGCCCTACCGCTTCTGCTCTGACCTCCACATGGA 392
 Db 103 ProArgMetIleGlyThrHisLeuThrValSerLeuLeuProLysSerPheTrpGluLys 122
 QY 393 GACTCCAGGTCTATATGCTGCAACCCCAAGGATCTGCTGCTTATTATCAG 452
 Db 123 LysCysLysTyrrValrValAlaArgAsnProLysAspValAlaValSerPheYrHis 142
 QY 453 TTCACCGCTCTCTCGGGACCATGAGCTACCGAGGACACCTTCAAGAATCTCGCGGAG 512
 Db 143 PheAspLysMetAsnGlnLeuHisProGluProGlyProTrpAspLysTyrrLeuGluLys 162
 QY 513 TTTATGAATGATAAGCTGGCTACGGCTCTCGTTTGAGCACGTCGAGGAGTCTGGGAG 572
 Db 163 PheMetGlnGlyLysValGlyTyrrGlyProTrpGlyProHisValArgAspTrpTrpGlu 182
 QY 573 CACCGCATGAGCTCGAAGCTGCTTTTCTCAAGTATGAGACATGCATCGGACCTGGTG 632
 Db 183 LeuArgLysLysGlnAsnMetLeuTyrrLeuPheTyrrGluAspMetIleGluAspProLys 202
 QY 633 ACGATGGTGAGCAGCTGCCAGATTCTCTGGGGGTCTCTGTGCAAGGCCCGAGCTGGA 692
 Db 203 ArgGluIleArgLysValIleSerPheLeuGlyLysAspLeuProGluThrIleValGlu 222
 QY 693 GCCCTGACGGACCATGCCAC-----CAGCTGGTGAGCAGCTGCTGCAACGCTGAG 743
 Db 223 LysIleCysGlnHisThrSerPheLysAlaMetLysGluAsnProLeuThrAsnTyrrSer 242
 QY 744 GCCCTGCC-----GTGGCGGGGAAGAGTT 770
 Db 243 SerValProSerAlaValMetAspGlnSerIleSerProPheMetArgLysGlyIleAla 262
 QY 771 GGGCTGTGGAAGGACATCTTACCGCTCTCCATGAATGAGAACTTTCAGCTTGTGTATAA 830
 Db 263 GlyAspTrpArgAsnHisPheThrGluAlaGlnSerGluArgPheAspLutTyrrGlu 282
 QY 831 CAGAAGTGGGAAAGTGACCTACGTTTGACTTT 866
 Db 283 GlyGluValAlaAlaThrAspLeuSerPheArgPhe 294

RESULT 81
 ID Q95JD6 CANFA PRELIMINARY; PRT; 295 AA.
 AC Q95JD6;
 DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
 DT 01-DEC-2001, sequence version 1.
 DT 07-FEB-2006, entry version 13.
 DE Sulfotransferase (Fragment).
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
 OC Canis.
 OX NCBI_TaxId=9615;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX TISSUE=Liver;
 RC MEDLINE=21285733; PubMed=11389699; DOI=10.1042/0264-6021.3560891;
 RA Tsoi C., Palany C.N., Morgenstern R., Swedmark S.;
 RT "Identification of a new subfamily of sulphotransferases: cloning and
 RT characterization of canine SUL1D1.";
 RL Biochem. J. 356:891-897(2001).
 CC
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC -----
 CC EMBL; AY004331; AAF86582.1; -; mRNA.
 DR HSP; P49888; 1HY3.

DR Ensembl; ENSAFG00000002873; Canis familiaris.
 DR GO; GO:008146; F:sulfotransferase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR InterPro; IPR000863; Sulfotransferase.
 DR Pfam; PF00685; Sulfotransferase 1; 1.
 DR ProDom; PD001218; Sulfotransferase; 1.
 DR Transferrase.
 KW NON_TER 295
 FT SEQUENCE 295 AA; 35207 MW; 62AA4C580B8CAF2B8 CRC64;
 SQ

Alignment Scores:
 Pred. No.: 2,596-20 Length: 295
 Score: 412.00 Matches: 89
 Percent Similarity: 53.7% Conservative: 65
 Best Local Similarity: 31.0% Mismatches: 115
 Query Match: 9.4% Indels: 18
 DB: 2 Gaps: 3

US-10-768-158-1 (1-2419) x Q95JD6_CANFA (1-295)
 QY 54 CCGGGGAGTTTCGAGAGCAAGTACTTTCAGATTCCATGGCGTGGCTGCGCCCTTCTGC 113
 Db 5 ProAspIleTyrrArgArgGluLeuValAspValGlnGlyValProLeuPheTrpSerIle 24
 QY 114 CCGGGGAAGATGGAGGAGATCGCCAACTTCGCGGTGCGGCCAGCAGCGTGGATCGTC 173
 Db 25 AlaGluGluTrpSerGlnValIleSerPheGluAlaArgProAspAspLeuLeuIleSer 44
 QY 174 ACCTACCCCAAGTCCGGCACCAGCTTGTGTCAGGAGGTGGTCTACTCTGGTAGCCAGGCG 233
 Db 45 ThrTyrrProLysSerGlyThrTrpValSerGluIleLeuAspLeuIleTyrrAsnAsn 64
 QY 234 GCTCACCCTGATGATCGGCTTGTATGAACATCCAGCAGCAGCTCCCGTCTCTGGAGTAC 293
 Db 65 GlyAspValGluLysCysLysArgAspAlaIleTyrrLysArgValProPheMetGluLeu 84
 QY 294 CCACAGCCG-----GGCTGGACATCATCAAGAACTGACCTCTCCCGCCCTC 341
 Db 85 IleIleProGlyPheGluAsnGlyIleGluAspLeuLysMetGlnProProArgLeu 104
 QY 342 ATCAAGACCCACCTGCGCTTCTGCGCTTCTGCGCTTCTGCGCTTCTGCGCTTCTGCGCT 401
 Db 105 ValLysThrHisLeuProValGlnLeuLeuProSerSerPheTrpLysAsnAsnCysLys 124
 QY 402 GTCATCTATATGCTCGCAACCCCAAGGATCTGCTGCTGCTTATTATCAGTTCCACCGC 461
 Db 125 MetValTyrrValAlaArgAsnAlaLysAspValAlaValSerTyrrTyrrPheTyrrGln 144
 QY 462 TCTCTCGGACCATGAGTACCGAGGACCTTTTCAAGAACTTCTGCGGAGGTTTATGAAT 521
 Db 145 MetAlaLysIleHisProLysAlaGlyThrTrpGluGluPheLeuAspLysPheMetThr 164
 QY 522 GATAAGCTGGCTACGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 581
 Db 165 GlyLysValAlaPheGlySerTrpTyrrAspHisValLysGlyTyrrTrpGluLysArgAsn 184
 QY 582 GACTCGAAGCTGCTTTTCTCAAGTATCAAGACATGCATCGGACCTCGGACCTGTCACGATG 641
 Db 185 AspTyrrArgIlePheTyrrLeuPheTyrrGluAspMetLysGluAsnProLysHisGluIle 204
 QY 642 GAGCAGCTGGCAGATTCCTGGGGGTGCTGTGTGACAAGGCCCGACCTCGAAGCCCTGACG 701
 Db 205 GlnLysLeuLeuGlnPheLeuLysAspLeuSerGluGluThrValAspLysIleLeu 224
 QY 702 GAGCAGCTGCCACGCTGCTG-----GACCAGTGTCTCAAGCTGAGGCGCTGAGGCGCTGCC 752
 Db 225 TyrHisSerSerPheAsnValMetLysGlnAsnProSerThrAsnTyrrThrThrIlePro 244
 QY 753 -----GTGGCGCGGGGAAGAGTTTGGGTGGTGG 779
 Db 245 AspPheAspMetAspHisSerValSerProPheMetArgLysGlyIleSerGlyAspTrp 264
 QY 780 AAGGACATCTTCCACGCTCTCCATGAATGAGAAGTTTGTGCTTGTATATAACAGAAGATG 839


```

RL  Genome Res. 10:1617-1630(2000).
RN  [7]
RP  STRAIN=C57BL/6J; TISSUE=Ovary and uterus;
RX  MEDLINE=2030913; PubMed=11076861; DOI=10.1101/gr.152600;
RA  Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA  Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA  Sumi N., Iehi Y., Nakamura S., Hazama M., Tashiro T., Harada A.,
RA  Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA  Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA  Onozaki Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA  Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT  "RIKEN integrated sequence analysis (RISA) system-384-format
RT  sequencing pipeline with 384 multicapillary sequencer.";
RL  Genome Res. 10:1757-1771(2000).
RN  [8]
RP  NUCLEOTIDE SEQUENCE.
RC  STRAIN=C57BL/6J; TISSUE=Ovary and uterus;
RA  Arahawa T., Carninci P., Fukuda S., Hashizume W., Hayashida K.,
RA  Hori F., Iida J., Imamura K., Imotani K., Itoh M., Kanagawa S.,
RA  Kawai J., Kojima M., Konno H., Murata M., Nakamura M., Ninomiya N.,
RA  Nishiyori H., Nomura K., Ohno M., Sakazume N., Sano H., Sasaki D.,
RA  Shibata K., Shiraki T., Tagami M., Tagami Y., Waki K., Watahiki A.,
RA  Muramatsu M., Hayashizaki Y.;
RL  Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
CC  Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC  Distributed under the Creative Commons Attribution-NoDerivs License
CC  -----
DR  EMBL; AK133530; BAE21709.1; -; mRNA.
DR  MGI; MGI:1926341; Sult1d1.
DR  GO; GO:0004062; F:aryl sulfotransferase activity; IDA.
DR  GO; GO:0000103; P:sulfate assimilation; IDA.
DR  InterPro; IPR000863; Sulfotransferase.
DR  Pfam; PF00685; Sulfotransfer 1; 1.
DR  ProDom; PD001218; Sulfotransferase; 1.
KW  Transferase.
SQ  SEQUENCE 295 AA; 35083 MW; A60A9AF60CC2736F CRC64;

Alignment Scores:
Pred. No.: 2,59e-20 Length: 295
Score: 412.00 Matches: 83
Percent Similarity: 55.7% Conservative: 63
Best Local Similarity: 31.7% Mismatches: 98
Query Match: 9.4% Indels: 18
DB: 2 Gaps: 2

US-10-768-158-1 (1-2419) x Q3UZZ6_MOUSE (1-295)
QY 129 GAGATCGCAACTTCCCGTGGCGCCGACGAGCTGTGGATCGTCACCTACCCCAAGTCC 188
Db 129 GAGATCGCAACTTCCCGTGGCGCCGACGAGCTGTGGATCGTCACCTACCCCAAGTCC 188
QY 30 GlnValGluSerPheGluAlaArgProAspAspIleLeuIleSerThrTyProLysSer 49
Db 30 GlnValGluSerPheGluAlaArgProAspAspIleLeuIleSerThrTyProLysSer 49
QY 189 GCACACAGCTTCTCGAGAGAGTGTCTACTTGTGTGAGCAGCGGCGCTGACCCCGATGAG 248
Db 189 GCACACAGCTTCTCGAGAGAGTGTCTACTTGTGTGAGCAGCGGCGCTGACCCCGATGAG 248
QY 50 GlyThrTrpValSerGluLeuLeuAspLeuIleTyAsnAsnGlyAspAlaGluLys 69
Db 50 GlyThrTrpValSerGluLeuLeuAspLeuIleTyAsnAsnGlyAspAlaGluLys 69
QY 249 ATCGCGTTCGATCAACATGACGAGCTCCCGGCTCGTGAGTACCACAGCGGCGCTG 308
Db 249 ATCGCGTTCGATCAACATGACGAGCTCCCGGCTCGTGAGTACCACAGCGGCGCTG 308
QY 70 CysLysArgAspAlaIleTyLysArgValProPheMetGluLeuIleProGlyIle 89
Db 70 CysLysArgAspAlaIleTyLysArgValProPheMetGluLeuIleProGlyIle 89
QY 309 -----GACATCATCAAGGAACCTGACCTCTCCCGGCTCATCAAGACCCACCTG 356
Db 309 -----GACATCATCAAGGAACCTGACCTCTCCCGGCTCATCAAGACCCACCTG 356
QY 357 CCTACCGCTTCTCGCTCTGACCTCCACATGAGACTCAAGTCACTATATGCT 416
Db 357 CCTACCGCTTCTCGCTCTGACCTCCACATGAGACTCAAGTCACTATATGCT 416
QY 110 ProValGlnLeuLeuProSerPheThrLysAsnAspCysLysIleIleTyValAla 129
Db 110 ProValGlnLeuLeuProSerPheThrLysAsnAspCysLysIleIleTyValAla 129
QY 417 CGCAACCCCAAGATCTGTGTGTCTTATTATCAGTTCACCGCTCTCTCGGACCATG 476
Db 417 CGCAACCCCAAGATCTGTGTGTCTTATTATCAGTTCACCGCTCTCTCGGACCATG 476
QY 130 ArgAsnAlaLysAspValValSerTyTyThrPheTyGlnMetAlaLysIleHis 149
Db 130 ArgAsnAlaLysAspValValSerTyTyThrPheTyGlnMetAlaLysIleHis 149

QY 477 AGCTACCGAGGACACCTTTCAAGAAATTCGCGGAGAGTTTATGAATGATAAGCTGGCTAC 536
Db 477 AGCTACCGAGGACACCTTTCAAGAAATTCGCGGAGAGTTTATGAATGATAAGCTGGCTAC 536
QY 150 ProGluProGlyThrTrpGluGluLeuGluLysPheMetAlaGlyGlnValSerPhe 169
Db 150 ProGluProGlyThrTrpGluGluLeuGluLysPheMetAlaGlyGlnValSerPhe 169
QY 537 GCCTCCTCGTTTGACAGCTGCGAGAGTCTCGGGAGACCGCATGACCTGCAACGTGCTT 596
Db 537 GCCTCCTCGTTTGACAGCTGCGAGAGTCTCGGGAGACCGCATGACCTGCAACGTGCTT 596
QY 170 GlyProTyrPheHisValLysSerTrpGluLysArgLysGluTyArgIleLeu 189
Db 170 GlyProTyrPheHisValLysSerTrpGluLysArgLysGluTyArgIleLeu 189
QY 597 TTTTCAAGTATGAAGACATGATCGGACCTGGTGGACGATGGTGAGCAGCTGGCCAGA 656
Db 597 TTTTCAAGTATGAAGACATGATCGGACCTGGTGGACGATGGTGAGCAGCTGGCCAGA 656
QY 190 TyrLeuPheTyGluAspMetLysGluAsnProLysCysGluIleGlnLysIleLeuLys 209
Db 190 TyrLeuPheTyGluAspMetLysGluAsnProLysCysGluIleGlnLysIleLeuLys 209
QY 657 TTTCTGGGGGTCTCTGTGTGACAGCCGAGCTGGAGAGCCCTGACGAGCAGCTGCCAC 713
Db 657 TTTCTGGGGGTCTCTGTGTGACAGCCGAGCTGGAGAGCCCTGACGAGCAGCTGCCAC 713
QY 210 PheLeuGluLysAspIleProGluGluIleLeuAsnLysIleLeuTyHisSerPhe 229
Db 210 PheLeuGluLysAspIleProGluGluIleLeuAsnLysIleLeuTyHisSerPhe 229
QY 714 -----CAGCTGGTGACGAGCTGCTGC 734
Db 714 -----CAGCTGGTGACGAGCTGCTGC 734
QY 230 SerValMetLysGluAsnProSerAlaSerTyThrMetMetLysGluGluMetAsp 249
Db 230 SerValMetLysGluAsnProSerAlaSerTyThrMetMetLysGluGluMetAsp 249
QY 735 AACGCTGAGGCGCTCCCGTGGCGGAGAGTGGGCTGTGGAAGGACATCTTCACC 794
Db 735 AACGCTGAGGCGCTCCCGTGGCGGAGAGTGGGCTGTGGAAGGACATCTTCACC 794
QY 250 HisSerValSerProPheMetArgLysGlyIleSerGlyAspTrpLysAsnGlnPheThr 269
Db 250 HisSerValSerProPheMetArgLysGlyIleSerGlyAspTrpLysAsnGlnPheThr 269
QY 795 GTCTCCATGAATGACAGATTTGACTTGTGTATTAACAGACAGATGGGAAGTGTGACCTC 854
Db 795 GTCTCCATGAATGACAGATTTGACTTGTGTATTAACAGACAGATGGGAAGTGTGACCTC 854
QY 270 ValAlaGlnTyGluLysPheGluGluAspTyValLysLysMetGluAspSerThrLeu 289
Db 270 ValAlaGlnTyGluLysPheGluGluAspTyValLysLysMetGluAspSerThrLeu 289
QY 855 ACGTTT 860
Db 290 LysPhe 291

RESULT 83
Q6NZD1_MOUSE PRELIMINARY; PRT; 295 AA.
AC Q6NZD1_MOUSE
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 13.
DE Sulfotransferase family 1D, member 1.
GN Name=Sult1d1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RX MEDLINE=22380257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalón D.K., Muzny K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W.,
RA Richardson D.K., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Greenwood J., Schmutz J., Myers R.M.,
RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.N., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Kidney;

```

```
RA Director MGC project;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBSJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BC066190; AAH66190.1; -; mRNA.
DR HSP; P49891; IAOU.
DR Ensembl; ENSMUSG00000029273; Mus musculus.
DR MGI; MGI:1926341; Sult1d1.
DR GO; GO:0004062; F:aryl sulfotransferase activity; IDA.
DR GO; GO:000103; P:sulfate assimilation; IDA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Transferase.
SQ SEQUENCE 295 AA; 35011 MW; A569PAP60CC0716F CRC64;

Alignment Scores:
Pred. No.: 2,59e-20 Length: 295
Score: 412.00 Matches: 83
Percent Similarity: 55.7% Conservative: 63
Best Local Similarity: 31.7% Mismatches: 98
Query Match: 9.4% Indels: 18
DB: 2 Gaps: 2

US-10-768-158-1 (1-2419) x QNZD1_MOUSE (1-295)
QY 129 GAGATCGCGCAATCTCCCGTGGCGCCAGCGAGTGTGATCGTCACCTACCCAGATCC 188
DB 30 GlnValGluSerPheGluAlaArgProAspIleLeuIleSerThrTyProLysSer 49
QY 189 GGCACGAGTCTGCTGAGAGTGTCTACTTGTGTGAGCCAGCGGCGTGCACCCGATGAG 248
DB 50 GlyThrThrTrpValSerGluIleLeuAspLeuIleTyAsnAsnGlyAspAlaGluLys 69
QY 249 ATCGGCTTGATGAACATGACGAGAGCTCCGGTCTCGAGTACCCAGCGCGGCTG 308
DB 70 CysLysArgAspAlaIleTyLysArgValProPheMetGluLeuIleLeuProGlyIle 89
QY 309 -----GACATCATCAAGAACTGACTCTCCCGGCTCTCAAGAGCCACCTG 356
DB 90 ThrAsnGlyValGluMetLeuAsnMetProSerProArgIleValLysThrHisLeu 109
QY 357 CCCTACCGCTTCTCCCTCTGACCTCCACATGAGAGTCCAGAGTCTATATAGCT 416
DB 110 ProValGlnLeuLeuProSerPheThrLysAsnAspCysLysIleIleTyValAla 129
QY 417 CGCAACCCCAAGATCTGGTGTGTCTATTATACAGTCCACCGCTCTCTGGGACCATG 476
DB 130 ArgAsnAlaLysAspValValSerTyTyTyTyTyTyTyTyTyTyTyTyTyTyTy 149
QY 477 AGCTACCGAGGACCTTTCAAGAATTCTCCCGAGGTTTATGAATGATAAGTGGGCTAC 536
DB 150 ProGluProGlyThrTrpGluGluPheLeuGluLysPheMetAlaGlyGlnValSerPhe 169
QY 537 GCCTCTGTTGAGCAGCTGAGAGTCTGGAGACCGCATCGATCGACCTGACAGCTGT 596
DB 170 GlyProTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTy 189
QY 597 TTTCTCAAGTATGAAGACATGATCGGACCTGGTGCAGTGTGTGAGAGCTGGCCAGA 656
DB 190 TyrLeuPheTyGluAspMetLysGluAsnProLysCysGluIleGlnLysIleLeuLys 209
QY 657 TTCTCGGGGTCTCTGTGACAGGCCACCTGGAAGCCCTGACGAGGACCTGCCAC--- 713
DB 210 PheLeuGluLysAspIleProGluIleLeuAsnLysIleLeuTyTyTyTyTyTyTy 229
QY 714 -----CAGCTGGTGGACAGTGTCTG 734
DB 230 SerValMetLysGlyAsnProSerAlaAsnTyThrThrMetMetLysGluGluMetAsp 249
QY 735 AACGCTGAGGCCCTGCCCGTGGCGGGGAGAGTGTGGCTGTGGAAGGACATCTTCACC 794

Db 250 HisSerValSerProPheMetArgLysGlyIleSerGlyAspTrpLysAsnGlnPheThr 269
QY 795 GTCTCCATGATGAGAGTTTGTGACTTGTGTATATAACAGAGATGGGANAAGTGTGACCTC 854
DB 270 ValAlaGlnTyGluLysPheGluGluAspTyValLysLysMetGluAspSerThrLeu 289
QY 855 ACGTTT 860
DB 290 LysPhe 291

RESULT 84
Q9Z1G0 RAT PRELIMINARY; PRT; 308 AA.
AC Q9Z1G0_
DT 01-MAY-1999, integrated into UniProtKB/TrEMBL.
DT 01-MAY-1999, sequence version 1.
DT 07-FEB-2006, entry version 20.
DE Tyrosine-ester sulfotransferase (EC 2.8.2.9).
GN Name=Sult1d1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Sprague-Dawley; TISSUE=Kidney;
RL Herrmann A., Stoffel W.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBSJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; U32372; AAC99890.1; -; mRNA.
DR HSP; P49888; 1HY3.
DR Ensembl; ENSRNOG0000001960; Rattus norvegicus.
DR RGD; 620491; Sult1d1.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0017067; F:tyrosine-ester sulfotransferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Transferase.
SQ SEQUENCE 308 AA; 36341 MW; 68719C9223BE9EAE CRC64;

Alignment Scores:
Pred. No.: 2,62e-20 Length: 308
Score: 412.00 Matches: 93
Percent Similarity: 52.8% Conservative: 59
Best Local Similarity: 32.3% Mismatches: 110
Query Match: 9.4% Indels: 26
DB: 2 Gaps: 5

US-10-768-158-1 (1-2419) x Q9Z1G0_RAT (1-308)
QY 63 TTCAGAGCAGTACTTCGAGTTCATGCGGTGGCTGCCGCCCTTCTGCGCGGGAAG 122
DB 8 PheArgArgGluLeuValAspValGlnGlyIleProLeuPheTrpSerIleAlaGluGln 27
QY 123 ATGGAGAGATCGCAACTTCCTCCGTGGCGCCAGCAGCTGTGGATCGTCACCTACCCC 182
DB 28 TrpSerGlnValGluSerPheGluAlaArgProAspIleLeuIleSerThrTyPro 47
QY 183 AAGTCCGACACAGCTGCTGACGAGTGTCTACTTGTGTGAGCCAGGCGCTGACCCC 242
DB 48 LysSerGlyThrThrTrpIleSerGluIleLeuAspLeuIleTyAsnAsnGlyAspAla 67
QY 243 GATGAGTCGGCTTCATGAACATGACGAGCTCCCGGTCCTGGAGTACCCACAGCCG 302
DB 68 GluLysCysLysArgAspAlaIleTyArgArgValProPheMetGluLeuIleLeu 87
QY 303 GGCCTG-----GACATCATCAAGAACTGACCTCTCCCGCCTCATCAAGAGC 350
```

```

Db 88 GlyIleThrAsnGlyValGluMetLeuAspAsnMetGlnSerProArgLeuValIleThr 107
      |||...:.....:..:|||||...:|||||...:|||||...:|||||...:
QY 351 CACCTGCCCTACCGCTTCTGCCCTCTGACCTCCACAAATGGAGACTCCAAGGTCTATCTAT 410
      |||...:..:|||||...:|||||...:|||||...:|||||...:|||||...:
Db 108 HisLeuProValGlnLeuLeuProSerSerPheTyrArgAsnAspCysLysMetIleTyr 127
      |||...:..:|||||...:|||||...:|||||...:|||||...:|||||...:
QY 411 ATGGCTCCGCAACCCAGAGATCTGGTGGTCTTATTATCAGTTCCACCGCTCTCGGG 470
      :|||...:..:|||||...:|||||...:|||||...:|||||...:|||||...:
Db 128 ValAlaArgAsnAlaLysAspValAlaValSerTyrTyrPheHisGlnMetAlaLys 147
      :|||...:..:|||||...:|||||...:|||||...:|||||...:|||||...:
QY 471 ACCATGAGCTACCGAGGACCTTCAAGAAATCTCGCGAGGTTTATGAATCAATAGCTG 530
      :|||...:..:|||||...:|||||...:|||||...:|||||...:|||||...:
Db 148 MetHisProGluProGlyThrTrpGluGluPheLeuGluLysPheMetAlaGlnVal 167
      :|||...:..:|||||...:|||||...:|||||...:|||||...:|||||...:
QY 531 GGCTACGGCTCTGTTTGGACAGCTGGAGGATTCTGGGACACCGCATGGAATCGAAC 590
      :|||...:..:|||||...:|||||...:|||||...:|||||...:|||||...:
Db 168 SerPheGlyProTyrAspHisValLysGlyTyrTrpGluLysArgLysGluTyrArg 187
      :|||...:..:|||||...:|||||...:|||||...:|||||...:|||||...:
QY 591 GTGCTTTTCTCAAGATGAAGACATGATCGGGACCTGGTGACGATGGTGGAGCAGCTG 650
      :|||...:..:|||||...:|||||...:|||||...:|||||...:|||||...:
Db 188 IleLeuTyrCysPheTyrGluAspMetLysGluAspProLysCysGluIleGlnLysVal 207
      :|||...:..:|||||...:|||||...:|||||...:|||||...:|||||...:
QY 651 GCCAGATTCTCG----- 662
      :|||...:..:|||||...:|||||...:|||||...:|||||...:|||||...:
Db 208 LeuLysPheLeuGluLysAspIleProGluGluValAlaAsnLysIleLeuTyrHisSer 227
      :|||...:..:|||||...:|||||...:|||||...:|||||...:|||||...:
QY 663 GGGGTGTCTGTGACAAAGCCAGCTGGAGGCC---CTGACGAGGACCTGCCACCAAGCTG 719
      :|||...:..:|||||...:|||||...:|||||...:|||||...:|||||...:
Db 228 SerPheSerValMetLysAlaAsnProSerAlaAsnTyrThrMetMetLysGluGlu 247
      :|||...:..:|||||...:|||||...:|||||...:|||||...:|||||...:
QY 720 GTGGACCAAGTGTGCAAGCTGAGGCCCTGGCCC---GTGGCCGGGGAGAGTTGGGCTG 776
      :|||...:..:|||||...:|||||...:|||||...:|||||...:|||||...:
Db 248 MetAspGlnSerValSer-----ProPheMetArgLysGlyIleSerGlyAsp 263
      :|||...:..:|||||...:|||||...:|||||...:|||||...:|||||...:
QY 777 TGGAGGACATCTTCACCGTCTCCATGAATGAGAAGTTTGACTTGGTGTATTAACAGAG 836
      :|||...:..:|||||...:|||||...:|||||...:|||||...:|||||...:
Db 264 TrpLysAsnGlnPheThrValAlaGlnTyrGluLysPheGluLysAspTyrValLysLys 283
      :|||...:..:|||||...:|||||...:|||||...:|||||...:|||||...:
QY 837 ATGGGAAAGTGTGACCTCAGCTTT 860
      :|||...:..:|||||...:|||||...:|||||...:|||||...:|||||...:
Db 284 MetGluGluSerThrLeuLysPhe 291
      :|||...:..:|||||...:|||||...:|||||...:|||||...:|||||...:

RESULT 85
ID Q9R2C2 MOUSE PRELIMINARY; PRT; 309 AA.
AC Q9R2C2
DT 01-MAY-2000, integrated into UniProtKB/TrEMBL.
DT 01-MAY-2000, sequence version 1.
DT 07-FEB-2006, entry version 21.
DE Tyrosine-ester sulfotransferase (EC 2.8.2.9).
GN Name=Sult1d1;
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=BALB/c; TISSUE=Kidney;
RA Herrmann A., Stoffel W.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DB EMBL; U32371; AAC99889.1; -; mRNA.
DR HSSP; P49888; 1HY3.
DR Ensembl; ENSMUSG0000029273; Mus musculus.
DR MGI; MGI:1926341; Sult1d1.
DR GO; GO:0004062; F:aryl sulfotransferase activity; IDA.
DR GO; GO:0000103; P:sulfate assimilation; IDA.
DR InterPro; IPR000863; Sulfotransferase.

```

```

DR Pfam; PF00685; Sulfotransfer_1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Transferase.
SQ SEQUENCE 309 AA; 36713 MW; 5D5E1CD3C6D5117 CRC64;

Alignment Scores:
Pred. No.: 2,62e-20 Length: 309
Score: 412.00 Matches: 83
Percent Similarity: 55.7% Conservative: 63
Best Local Similarity: 31.7% Mismatches: 98
Query Match: 9.4% Indels: 18
DB: 2 Gaps: 2

US-10-768-158-1 (1-2419) x Q9R2C2_MOUSE (1-309)

QY 129 GAGATCGCAACTTCCCGTGGCCGACGAGCTGTGGATCGTCACCTACCCCAAGTCC 188
      :|||...:..:|||||...:|||||...:|||||...:|||||...:|||||...:
Db 30 GlnValGluSerPheGluAlaArgProAspAspIleLeuIleSerThrTyrProLysSer 49
      :|||...:..:|||||...:|||||...:|||||...:|||||...:|||||...:
QY 189 GGCACCAGCTTGTGTCAGGAGGTGTCTACTTGTGACCGAGGCGCTGACCCCATGAG 248
      :|||...:..:|||||...:|||||...:|||||...:|||||...:|||||...:
Db 50 GlyThrThrTrpValSerGluIleLeuAspLeuIleTyrAsnAsnGlyAspAlaGluLys 69
      :|||...:..:|||||...:|||||...:|||||...:|||||...:|||||...:
QY 249 ATCGGCTTGATGAACATCGACGAGCTCCCGGTCTGGAGTACCCACAGCCGGGCTG 308
      :|||...:..:|||||...:|||||...:|||||...:|||||...:|||||...:
Db 70 CysLysArgAspAlaIleTyrLysArgValProPheMetGluLeuIleIleProGlyIle 89
      :|||...:..:|||||...:|||||...:|||||...:|||||...:|||||...:
QY 309 -----GACATCATCAAGAACTGACCTCTCCCGCTCATCAAGAGCCACCTG 356
      :|||...:..:|||||...:|||||...:|||||...:|||||...:|||||...:
Db 90 ThrAsnGlyValGluMetLeuAsnAsnMetProSerProArgIleValLysThrHisLeu 109
      :|||...:..:|||||...:|||||...:|||||...:|||||...:|||||...:
QY 357 CCTTACCGCTTCTGCGCTCTGACCTCCACAAATGGAGACTCCCAAGTCAATATATGGCT 416
      :|||...:..:|||||...:|||||...:|||||...:|||||...:|||||...:
Db 110 ProValGlnLeuLeuProSerSerPheTrpLysAsnAspCysLysIleIleTyrValAla 129
      :|||...:..:|||||...:|||||...:|||||...:|||||...:|||||...:
QY 417 CGCAACCCCAAGGATCTGGTGTCTTATTTATCAGTTCCACCGCTCTCTCGGACCATG 476
      :|||...:..:|||||...:|||||...:|||||...:|||||...:|||||...:
Db 130 ArgAsnAlaLysAspValValSerTyrTyrPheTyrGlnMetAlaLysIleHis 149
      :|||...:..:|||||...:|||||...:|||||...:|||||...:|||||...:
QY 477 AGCTACCGAGGACCTTTCAAGAATTTCTGCGGAGTTTATGAATGATAAGCTGGGCTAC 536
      :|||...:..:|||||...:|||||...:|||||...:|||||...:|||||...:
Db 150 ProGluProGlyThrTrpGluGluPheLeuGluLysPheMetAlaGlyGlnValSerPhe 169
      :|||...:..:|||||...:|||||...:|||||...:|||||...:|||||...:
QY 537 GGTCTCTGGTTTGACACGTGACGAGTTCTGGGAGACCGCATGGACTCGAACCTGTCT 596
      :|||...:..:|||||...:|||||...:|||||...:|||||...:|||||...:
Db 170 GlyProTyrAspHisValLysSerTrpTrpGluLysArgLysGluTyrArgIleLeu 189
      :|||...:..:|||||...:|||||...:|||||...:|||||...:|||||...:
QY 597 TTTCTCAAGTATGAACATCGATCGGACCTGGTGACGATGGTGGAGCAGCTGCCACA 656
      :|||...:..:|||||...:|||||...:|||||...:|||||...:|||||...:
Db 190 TyrLeuPheTyrGluAspMetLysGluAsnProLysCysGluIleGlnLysIleLys 209
      :|||...:..:|||||...:|||||...:|||||...:|||||...:|||||...:
QY 657 TTCTCTGGGGTGTCTGTGACAAAGCCAGCTGGAAAGCCCTGACGGAGCACTGCCAC 713
      :|||...:..:|||||...:|||||...:|||||...:|||||...:|||||...:
Db 210 PheLeuGluLysAspIleProGluGluLeuAsnLysIleLeuTyrHisSerPhe 229
      :|||...:..:|||||...:|||||...:|||||...:|||||...:|||||...:
QY 714 -----CAGCTGGTGGACAAGTGTGCTGC 734
      :|||...:..:|||||...:|||||...:|||||...:|||||...:|||||...:
Db 230 SerValMetLysGluAsnProSerAlaAsnTyrThrMetLysGluGluMetAsp 249
      :|||...:..:|||||...:|||||...:|||||...:|||||...:|||||...:
QY 735 AACGCTAGGCGCTGCGCGTGGCGGAGAGAGTTGGCTGTGGAGGACATCTTCACC 794
      :|||...:..:|||||...:|||||...:|||||...:|||||...:|||||...:
Db 250 HisSerValSerProPheMetArgLysGlyIleSerGlyAspTrpLysAsnGlnPheThr 269
      :|||...:..:|||||...:|||||...:|||||...:|||||...:|||||...:
QY 795 GTCTCCATGAATGAGAGTTTGACTTGGTGTATTAACAGAGATGGGAAAGTGTGACCTC 854
      :|||...:..:|||||...:|||||...:|||||...:|||||...:|||||...:
Db 270 ValAlaGlnTyrGluLysPheGluLysPyrValLysLysMetGluAspSerThrLeu 289
      :|||...:..:|||||...:|||||...:|||||...:|||||...:|||||...:
QY 855 ACCTTT 860
      :|||...:..:|||||...:|||||...:|||||...:|||||...:|||||...:
Db 290 LysPhe 291
      :|||...:..:|||||...:|||||...:|||||...:|||||...:|||||...:

RESULT 86
Q66KW4_XENLA

```


Db 9 LeuAlaGluGluMetGluAsnPheGlnValThrMetGlyHisIleGlu----- 24
 Qy 81 GAGTTCCATCGCGTGGCTGCCG----CCTTCTGCCGCGGGAAGATGAGGAGATCGCC 137
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 25 -----GlyValProLeuProGlyProThrCys--AspGluTrpAspThrIleTyr 40
 Qy 138 AACTTCGGGTGGCGGCCACGACGTGTGGATGTGCACCTACCCCAAAGTCCGCCACCAGC 197
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 41 AsnPheGlnAlaArgGluAspAspIleLeuIleAlaThrTyrProLysSerGlyThrThr 60
 Qy 198 TTGCTCACGAGGTGTCTACTTGTTGGTGCAGCCAGGGCGCTGCACCCCATGAGATCGCGCTTG 257
 :::||:||||:||||:||||:||||:||||:||||:||||:||||:
 Db 61 TrpMetGlnGluIleValAspLeuIleLeuGlnGluGlyAspValGlnLysSerMetArg 80
 Qy 258 ATGAACATGCACGACGAGCTCCCGTCTCGGATGATCCACAGCGC----- 302
 ::::||||:||||:||||:||||:||||:||||:||||:||||:
 Db 81 AlaProCysPheIleLysValProPheIleGluMetIleProProLysSerMetProSer 100
 Qy 303 GGCTTGGACATCAAGAAGTACCTCTCCCGCTCATCAAGAGCCACCTGCGCTTAC 362
 |||||:||||:||||:||||:||||:||||:||||:||||:||||:
 Db 101 GlyLeuGluLeuAlaLysThrMetLysSerProArgIleLeuLysThrHisLeuProIle 120
 Qy 363 CGCTTTCTGCCCTCTACCTCCACATGAGACTCCAAGTCACTCATATATGGCTGCCAAC 422
 ||||| ||||| :|||:||||:||||:||||:||||:||||:||||:
 Db 121 AsnLeuLeuProProSerPheTrpGluLysAsnAlaLysValValYrValAlaArgAsn 140
 Qy 423 CCCAAGATCTGGTGTCTATTATTCAGTTCCACCGCTCTCTGCGGACCATGAGCTAC 482
 ||||| :|||:||||:||||:||||:||||:||||:||||:||||:
 Db 141 AlaLysAspCysMetValSerTyrTyrPheHisLysMetAsnThrPheLeuLeuAsp 160
 Qy 483 CGAGGCACCTTCAAGAAATCTGCCGAGGTTTATGAATGATTAAGTGGCTTACGGCTCC 542
 |||||:||||:||||:||||:||||:||||:||||:||||:||||:
 Db 161 ProGlyThrTrpAspAsnPhePheSerGluPheLeuSerGlyAspValProTrpGlySer 180
 Qy 543 TGTGTTGACACGCTGAGAGTTCTGGGACGCCGATGAGCTCGAACGTGCTTTTCTC 602
 |||||:||||:||||:||||:||||:||||:||||:||||:||||:
 Db 181 TrpPheAspHisValLeuGlyTyrTrpTrpLysAlaMetAspLysHisGlnIleLeuPheIle 200
 Qy 603 AACTATGAACATGATCGGACCTCGTGACGATGGTGGAGCAGCTGGCCAGATTCCTG 662
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 201 PheTyrGluAspMetIleGluAspProMetArgGluIleArgLysValMetLysPheLeu 220
 Qy 663 GGGGTGCTCTGTGACAAGGCCAGCTGCAAGCCCTGACGGAGCATCGCCAC----- 713
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 221 GlyLysAspLeuSerAspGluAlaLeuGluAsnValLysTyrHisSerPheGlnAla 240
 Qy 714 CAGCTGGTGACACGCTGTCACCGCTGAGCGCTGAGCGCTGCC----- 752
 :||: ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 241 MetLysGluAsnProMetThrAsnAsnSerThrValProAsnSerIleMetAspAspThr 260
 Qy 753 -----GTGGGCGGGGAAGATGTGGGCTGTGGAAGGACATCTTTCACCGCTCC 800
 ::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 261 IleSerProPheMetArgLysGlyIleValGlyAspTrpLysThrHisPheSerValThr 280
 Qy 801 ATGAATCAGAAGTTTGCATTTGTTGTATAACAGAGAAGTGGAAAAGTGTGACCTCACTTT 860
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 281 GlnAsnPheIlePheAspLysGluTyrLysLysLysMetGluGlySerGlyLeuAsnPhe 300

RESULT 87
 QSHZUO_XENTR
 ID QSHZUO_XENTR PRELIMINARY; PRT; 297 AA.
 AC QSHZUO;
 DT 15-FEB-2005, integrated into UniProtKB/TrEMBL.
 DT 15-FEB-2005, sequence version 1.
 DT 07-FEB-2006, entry version 10.
 DE Hypothetical LOC496998.
 GN Name=LOC496998;
 OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 OC Xenopodinae; Xenopus; Silurana.
 OX NCBI_TaxID=8364;
 RN [1]

RP NUCLEOTIDE SEQUENCE.
 RX TISSUE=Whole body;
 RA MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.2426303899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny K.B., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smalish D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Whole body;
 RA Klein S., Gerhart D.S.;
 RA Submitted (JAN-2005) to the EMBL/GenBank/DBJ databases.
 CC
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC
 DR EMBL; BC088888; AAH88888.1; -; mRNA.
 DR GO; GO:0008146; F:sulfotransferase activity; IEA.
 DR InterPro; IPR000863; Sulfotransferase.
 DR Pfam; PF00685; Sulfotransferase 1; 1
 DR ProDom; PD001218; Sulfotransferase; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 297 AA; 34863 MW; E8248162DC31D7D2 CRC64;
 Alignment Scores:
 Pred No.: 3,31e-20 Length: 297
 Score: 410.50 Matches: 85
 Percent Similarity: 55.7% Conservative: 61
 Best Local Similarity: 32.4% Mismatches: 99
 Query Match: 9.3% Indels: 17
 DB: 2 Gaps: 4
 US-10-768-158-1 (1-2419) x Q5HZU0_XENTR (1-297)
 Qy 120 AAGTGGAGAGATGCCCACTCCCGTGGCGGCCCGCCAGCAGCTGTGGTGGTCACTAC 179
 Db 30 LysileAspSerIleGlnAspPheLysValLysAspThrAspValPheLeuValThrTyr 49
 Qy 180 CCAAGTCCGACGAGCTTCTGTCAGGAGGTGTCTACTTGGTG---AGCCAGGCGCT 236
 Db 50 ProlysThrGlyThrIleTrpThrGlnGlnLeuSerLeuIlePheAsnGluGlyHis 69
 Qy 237 GACCCGATGATCGGCTTGATGACATCGACGAGCTCCCGTCTCGAGTACCCA 296
 Db 70 ArgAsnGlyThrGluAlaIleAlaAsnVal---PheArgValProTrpIleGluThr 88
 Qy 297 CAGCGCGCCTGGACATCATCAAGAACTGACCTCTCCCGCTCATCAAGACCCACCTG 356
 Db 89 HisSerLysValAsp---TyrAspSerArgProSerProArgLeuPheSerSerHisLeu 107
 Qy 357 CCTTACCGCTTCTGCTCTGACCTCCACATGAGAGCTCAAGGTCACTATATGCT 416
 Db 108 ProHisThrLeuValProLysAspLeuArgAsnLysLysGlyLysIleIleTyrValGly 127
 Qy 417 CGCACCCCAAGGATCTGGTGTCTTATATCATGTTCCACCGCTCTCTGGACCATG 476
 Db 128 ArgAsnProLysAspAlaAlaValSerTyrHisPheTyrAsnValIleValArgLeu 147

Qy 477 AGCTACCGAGGACCTTTCAAGAAATTCTGCCGAGAGTTTATGAATGATAAGCTGGGCTAC 536
 Db 148 LysGlnValAsnAspTrpGluSerPheLeuAspArgTyrLeuThrGlyGluValLeuGly 167
 Qy 537 GGCTCTCTGGTTTGGACGCTGTCAGGAGTTCTGGGAGCACCAGCATCGACTCGAACCTGCTT 596
 Db 168 GlySerTrpPheAspHisValLysGlyTyrThrHisGlnGluAspPheAsnIleLeu 187
 Qy 597 TTTCTCAAGTATGAAGACATCATCGGACCTGTGTGACGATGGTGGAGCAGCTGGCCAGA 656
 Db 188 PheValThrTyrGluGluMetLysLysAspLeuArgSerAlaValLeuLysIleCysLys 207
 Qy 657 TTCCTGGGGGTGTCTGTGACCAAGCCCGCTGGAAGCCCTGACCGAG----- 704
 Db 208 PheValGluLysGluLeuAsnGluGlnValAspThrIleValGluLysAlaThrPhe 227
 Qy 705 -----CACTGCCACAGCTGTGTGGACCGAGTGTGCTGCTGC 734
 Db 228 LysAsnMetLysHisAspProLeuAlaAsnTyrThrAsnValSerThrAspHisLeuAsp 247
 Qy 735 AACCTGAGGCCCTGCCCGCGGCGGAGAGTTGGCTGTGGAAGACATCTTCACC 794
 Db 248 MetLysAsnGlyThrPheLeuArgGlyThrValGlyAspTrpLysLeuMetThr 267
 Qy 795 GTCTCCATGAATGAGAGTTTGTGCTGTATAAACAGAGATGGGAAAGTGTGACCTC 854
 Db 268 ValAlaGlnAsnGluLysPheAspLysIleTyrSerGluLysMetLysGlyValProIle 287
 Qy 855 ACGTTT 860
 Db 288 AsnPhe 289
 RESULT 88
 ID ST1E1_CAVPO STANDARD; PRT; 296 AA.
 AC P49887;
 DT 01-OCT-1996, integrated into UniProtKB/Swiss-Prot.
 DT 01-OCT-1996, sequence version 1.
 DT 07-FEB-2006, entry version 43.
 DE Estrogen sulfotransferase (EC 2.8.2.4) (Sulfotransferase, estrogen-
 DE preferring) (ST1E3).
 GN Name=ST1E1; Synonyms=STE;
 OS Cavia porcellus (Guinea pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 OC Hystricognathi; Caviidae; Cavia.
 OC NCBI_TaxID=10141;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [MRNA], AND PARTIAL PROTEIN SEQUENCE.
 RC STRAIN=NIH 2; TISSUE=Adrenal cortex;
 RX MEDLINE=93024479; PubMed=1406700; DOI=10.1210/me.6.8.1216;
 RA Oeda T., Lee Y.C., Driscoll W.J., Chen H.-C., Strott C.A.;
 RT "Molecular cloning and expression of a full-length complementary DNA
 RT encoding the guinea pig adrenocortical estrogen sulfotransferase";
 RL Mol. Endocrinol. 6:1216-1226(1992).
 RN [2]
 RP PAPS-BINDING SITE, AND MUTAGENESIS.
 RX MEDLINE=95071451; PubMed=7980593;
 RA Komatsu K., Driscoll W.J., Koh Y., Strott C.A.;
 RT "A P-loop related motif (GxxGxxK) highly conserved in
 RT sulfotransferases is required for binding the activated sulfate
 RT donor";
 RL Biochem. Biophys. Res. Commun. 204:1178-1185(1994).
 CC -!- FUNCTION: Sulfation of estradiol and estrone. May control the
 CC level of the estrogen receptor by sulfurylating free estradiol.
 CC -!- CATALYTIC ACTIVITY: 3'-phosphoadenyl sulfate + estrone =
 CC adenosine 3',5'-bisphosphate + estrone 3-sulfate.
 CC -!- SUBUNIT: Homodimer.
 CC -!- SUBCELLULAR LOCATION: Cytoplasm.
 CC -!- TISSUE SPECIFICITY: Adrenal gland and much less in liver. Uterine
 CC STE is detectable only during pregnancy.
 CC -!- PTM: The N-terminus is blocked.

CC -1- SIMILARITY: Belongs to the sulfotransferase family.
CC
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC
DR EMBL; U09552; AAA18495.1; -; mRNA.
DR PIR; A44011; A44011.
DR HSP; P49888; 1G3M.
DR SMR; P49887; 5-234.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Direct protein sequencing; Lipid-binding; Steroid-binding;
KW Transferrase.
FT CHAIN 1 296 Estrogen sulfotransferase.
FT /FTID=PRO_0000085152.
FT NP_BIND 49 54 PAPS (By similarity).
FT NP_BIND 131 139 PAPS (By similarity).
FT NP_BIND 194 230 PAPS (By similarity).
FT NP_BIND 258 260 PAPS (By similarity).
FT ACT_SITE 109 109 Proton acceptor (By similarity).
FT MUTAGEN 260 260 G->A: Loss of activity.
FT MUTAGEN 263 263 G->A: Loss of activity.
FT MUTAGEN 266 266 K->A: Loss of activity.
SQ SEQUENCE 296 AA; 35245 MW; 019D92895E29901D CRC64;

Alignment Scores:
Pred. No.: 4,598-20 Length: 296
Score: 408.50 Matches: 94
Percent Similarity: 52.4% Conservative: 50
Best Local Similarity: 34.2% Mismatches: 112
Query Match: 9.3% Indels: 19
DB: 1 Gaps: 4

US-10-768-158-1 (1-2419) x ST1E1_CAVPO (1-296)

QY 72 AAGTACTTC---GAGTTCATCGCTGGCGCTGCCCGCTTCGCGGGAGATGAG 128
DB 11 GlutyrPheAspGluPheArgGlyIleLeuLeuTyrLysGlnPheIleLysTrpAsp 30
QY 129 GAGATCGCAACTTCCTCCGCTGGCGCCGACGACGTGTGGATCGTCACCTACCCCAAGTCC 188
DB 31 AenValGluAlaPheGlnAlaArgProAspLeuValIleAlaAlaTyrProLysSer 50
QY 189 GGCACAGCTTCCTCAGAGGTGTCTACTTGTGTAGCCGGCGCTGACCCCATGAG 248
DB 51 GlyThrThrTrpIleSerGluValValCysMetIleTyrAlaGluGlyAspValLysLys 70
QY 249 ATCGCTTGTATGAACATCAGCAGCAGCTCCGCTCGGTGAGTACCCACAGCCG----- 302
DB 71 CysArgGlnAspAlaIlePheAsnArgValProPheLeuGluCysArgAsnAspLysMet 90
QY 303 -----GGCTGGACATCATCAAGAACTGACCTTCCTCCGCTCATCAAGAGCCACCTG 356
DB 91 MetAenGlyValLysGlnLeuGluMetAenSerProArgIleIleLysThrHisLeu 110
QY 357 CCCTACCGCTTCTCCCTCTGACCTCCACATGAGACTCCAAAGTATCATATATGCT 416
DB 111 ProProArgLeuLeuProAlaSerPheTrpGluLysArgCysLeuMetIleCysLeu 130
QY 417 GCGACCCCAAGGATCTGGTGTCTATTATCATGTTCCACCGCTCTCTCGGGACCATG 476
DB 131 ArgAsnAlaLysAspValAlaValSerTyrTyrPhePheLeuMetValAlaHis 150
QY 477 AGCTACCGAGGACCTTTCAAGAACTTCGCGGAGGTTTATGAATCAATGAAGCTGGCTAC 536
DB 151 ProAspProGlySerPheProGluPheValGluLysPheMetGlnGlyLysValProTyr 170
QY 537 GCCTCTGTTGAGCAGCTGAGGAGTCTGGGAGCAGCCGCTGACTCGAACCTGCTT 596
DB 171 GlySerTrpTyrAspHisValLysSerTrpTrpGluLysSerThrAspProArgIleLeu 190
QY 597 TTTCTCAAGTATGAAGATCATGATCGGAGCTGGTGGTGGACGATGCTGGAGCAGCTGGCCAGA 656

Db 191 PheIlePheTyrGluAspMetLysGluAspIleArgLysGluValLeuLeuHis 210
QY 657 TTCTTGGGGTGTCTGTGACAGCCAGCTGGAAGCCCTGAGGAGCAGCTGCCAC--- 713
Db 211 PheLeuGlyArgLysProSerGluGluLeuValAspLysIleIleLysHisThrSerPhe 230
QY 714 -----CAGCTGGTGGACCACTGCTGCCAAC 737
Db 231 GlnGluMetLysAsnAsnProSerThrAsnTyrThrMetLeuProGluLeuMetAsn 250
QY 738 GCTGAGGCCCTTGGCC---GTGGCGGGGAGAGCTGTGGCTGTGGAGGACATCTTACC 794
Db 251 GlnLysValSerProPheMetArgLysGlyIleSerGlyAspTrpLysAsnHisPheThr 270
QY 795 GTCTCCATGAATGAGAGTTTGACTTGGTGTATTAACAGAGATG 839
Db 271 ValAlaLeuAsnGluSerPheAspLysHisTyrGlnGlnMet 285

RESULT 89
O70262 MOUSE PRELIMINARY; PRT; 304 AA.
AC O70262;
DT 01-AUG-1998, integrated into UniProtKB/TrEMBL.
DT 01-AUG-1998, sequence version 1.
DT 07-FEB-2006, entry version 26.
DE Phenol sulfotransferase.
GN Name=Sult1c1; Synonyms=Sult1a2;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ddv; TISSUE=Olfactory;
RX MEDLINE=96208491; PubMed=8641270;
RA Miyawaki A., Homma H., Tamura H., Matsui M., Mikoshiba K.;
RT "Zonal distribution of sulfotransferase for phenol in olfactory
sustentacular cells.";
RL EMBO J. 15:2050-2055(1996).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ddv; TISSUE=Olfactory;
RX MEDLINE=9828237; PubMed=9560327;
RA Tamura H.O., Harada Y., Miyawaki A., Mikoshiba K., Matsui M.;
RT "Molecular cloning and expression of a cDNA encoding an olfactory-
specific mouse phenol sulphotransferase.";
RL Biochem. J. 331:953-958(1998).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ddv; TISSUE=Olfactory;
RA Tamura H.-O., Harada Y., Miyawaki A., Mikoshiba K., Matsui M.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
CC
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC
DR EMBL; AF033653; AAC17740.1; -; mRNA.
DR HSP; P50224; 1CJM.
DR MGI; MGI:102928; Sult1c1.
DR GO; GO:0005737; Cytoplasm; IDA.
DR GO; GO:0004062; F-aryl sulfotransferase activity; IDA.
DR GO; GO:0006790; P:sulfur metabolism; IDA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Transferrase.
SQ SEQUENCE 304 AA; 35797 MW; 30857261A1462BE4 CRC64;

Alignment Scores:
Pred. No.: 4,628-20 Length: 304


```
QY 375 TCTGACCTCCAAATGGAGACTCCAAAGTTCATATATGGCTCGCAACCCCAAGATCTG 434
DQ 130 ProAlaPheTrpGluAsnLysCysLysThrIleTyrValAlaAArgAsnAlaLysAspAen 149
QY 435 GTGGTGCTTATTATTCAGTTCACCGCTCTCTCGGACCATGAGTACTACGAGGACACCTTT 494
DQ 150 LeuValSerTyrTyrPhePheAspCysMetAenMetThrGlnProGluProGlyThrMet 169
QY 495 CAAGAATCTCGCGAGGATTTATGAATGATAGCTGGGCTACGGCTCTGCTGTGAGCAGC 554
DQ 170 GluGluTyrIleHisLysPheMetArgGlyGluLeuSerTrpGlySerTrpTyrAspHis 189
QY 555 GTGAGGAGTCTGGGACACCGCATGACGTCGAAGCTGCTTTCTCACTATGATGAC 614
DQ 190 ValLysGlyTyrTrpLysGluLysAspAsnLysAsnIleLeuTyrLeuPheTyrGluAsp 209
QY 615 ATGCATCGGACCTGGTACGATGCTGGAGCAGCTGGCAGATCTCTGGGGTGTCTCTGT 674
DQ 210 MetLysGluAenProArgGluValGluArgIleMetArgTyrLeuAspValSerVal 229
QY 675 GACAAAG-----GCCAGCTGGAAGCCCTGACGAGGACACTGCCACAGCTGGTGAC 725
DQ 230 SerAspGluValIleSerLysIleValGluLeuThrSerPheGluLysMetLysGluAen 249
QY 726 CAGTGCTGCAAGCTGAGGCCCTGCC----- 752
DQ 250 ProMetAlaAenTyrThrCysValProAlaProValPheAspHisSerLysSerProPhe 269
QY 753 GTGGGCCCGGGAAGAGTGTGGGTGTGGAAGGACATCTTCACCGTCTCCATGAATGAGAAG 812
DQ 270 MetArgLysGlyLysValGlyAspTrpArgAenTyrPheThrProGluGlnGluLysMet 289
QY 813 TTGTGACTGGTGATATAACAGAAAGATGGAAAGTGTGACCTCACCCTTT 860
DQ 290 PheGluGluAspTyrLysGluGlnMetLysAspValAspIleProPhe 305
RESULT 91
Q3VHH0 DROME PRELIMINARY; PRT; 316 AA.
AC Q3VHH0;
DT 01-MAY-2000, integrated into UniProtKB/TrEMBL.
DT 01-MAY-2000, sequence version 1.
DT 07-FEB-2006, entry version 27.
DE CG16733-PA.
GN ORFNames=CG16733, Dmel_CG16733;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Cochran S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieri S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwa C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
```

```
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamis I., Simpson M., Skupski M.P., Smith T.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195 (2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Wheeler S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: release 3 of the Drosophila
melanogaster euchromatic genome sequence.";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079 (2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
a genomics perspective.";
RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20 (2002).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Battencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
systematic review.";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22 (2002).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RX Berkeley Drosophila Genome Project;
RA Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
RA Hoskins R., Stapleton M., Pacleb J., Park S., Svirskas R., Smith E.,
RA Yu C., Rubin G.;
RT "Drosophila melanogaster release 4 sequence.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [6]
RP NUCLEOTIDE SEQUENCE.
RX FlyBase;
RL Submitted (JAN-2006) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
EMBL; AE003682; AAF54344.1; --; Genomic_DNA.
HSSP; P49888; IHY3.
FlyBase; FBgn0037665; CG16733.
GO; GO:0008146; F:sulfotransferase activity; IBA.
InterPro; IPR000863; Sulfotransferase.
```


QY 366 TTTTCTGCTCTGACCTCCACAAATGAGAGACTCAAGGTCACTATATATGCTCGCAACCCC 425
DB 104 IIPheProLysAlaPheTherSerLysAlaLysIleIleTyrValMetArgAsnPro 123
QY 426 AAGGATCTGGTGGTCTTATTATACAGTTCCACCGCTCTCTCGGAGACCATGAGCTACCGA 485
DB 124 LysAspIlePheValSerLeuPheTyrPheAlaLysIleIleCysHisTyrLysAspPro 143
QY 486 GGCACCTTTCAAGAAATTCGCGGAGGTTTATGAATGATAAGCTGGGCTACGGCTCCTGG 545
DB 144 GluSerPheGluGlnTyrLeuGluAspPheLeuGlnGlyAsnIleLeuTyrAsnSerTrp 163
QY 546 TTTGAGCAGCTGAGAGTCTGGGAGACCGCATGGACTCGAAGCTGCTTTTCTCTCAAG 605
DB 164 PheGluHisValArgGlyTrpMetGlnMetLysAspAsnSerAsnPhePheIleIleThr 183
QY 606 TATGAAGACATCATCGGACCTGGTGACGATGGTGAGCAGCTGGCCAGATTCTCTGGGG 665
DB 184 TyrGluGluLeuHisGlnAspLeuArgGlyCysIleLeuArgIleCysLysPheLeuGly 203
QY 666 GTGTCTGTGACAAGCCAGCTGGAAGCCCTGACGGAGCAC----- 707
DB 204 LysGluLeuGlyAspSerLysIleAspThrIleValLysHisSerSerPheGluAlaMet 223
QY 708 -----TGCCACACGCTGGTGACGAGTyrLeuGlyThrGluGluTyrIleAsp----- 240
DB 224 LysGluAsnLysMetSerAsnTyrSerLeuGlyThrGluGluTyrIleAsp----- 240
QY 735 AAGCTGAGCGCCCTGCCCTGGCGGGGAGAGTTGGCTGTGAAGGACATCTTCACC 794
DB 241 HisThrLysGlyThrPheMetArgLysGlyThrAlaGlyAspTrpLysAsnHisPheThr 260
QY 795 GTCTCCATGAATGAGAAGTTTGACTGTGGTGTATAACAGAAAGTGTGACCTC 854
DB 261 ValAlaGlnSerGluHisPheAspArgValTyrGlnGluLysMetLys-----AspLeu 278
QY 855 ACGTTTGACTTTTAT 869
DB 279 AsnMetLysPhePhe 283
RESULT 94
ID Q491K6_BRARE PRELIMINARY; PRT; 293 AA.
AC Q491K6;
DC 13-SEP-2005, integrated into UniProtKB/TrEMBL.
DT 13-SEP-2005, sequence version 1.
DT 07-FEB-2006, entry version 2.
DE SULT1 isoform 5.
GN Name=sult1s5;
OS Brachydanio rerio (zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed16045754; DOI=10.1111/j.1742-4658.2005.04791.x;
RA Yasuda S., Kumar A.P., Liu M.Y., Sakakibara Y., Suiko M., Chen L.,
RA Liu M.C.;
RT "Identification of a novel thyroid hormone-sulfating cytosolic
RT sulfotransferase, SULT1 S5, from zebrafish.";
RL FEBS J. 272:3828-3837(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Liu M.-C.;
RL Submitted (JAN-2005) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; AY879099; AAY47051.1; -; mRNA.
DB ZFIN; ZDB-GENE-050809-2; sult1s5.

DR GO; GO:0008146; F:sulfotransferase activity; IDA.
DR GO; GO:0006805; P:xenobiotic metabolism; IDA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase_1.
DR ProDom; PD001218; Sulfotransferase; 1.
SQ SEQUENCE 293 AA; 34454 MW; 108BB08C28C4A3BC CRC64;
Alignment Scores:
Pred. No.: 1.44e-19 Length: 293
Score: 401.50 Matches: 93
Percent Similarity: 48.1% Conservative: 44
Best Local Similarity: 32.6% Mismatches: 121
Query Match: 9.1% Indels: 27
DB: 2 Gaps: 4
US-10-768-158-1 (1-2419) x Q491K6_BRARE (1-293)
QY 81 GAGTTCATGCGTGGCTGGCTGCGGCTCTCTCGCGGGAAGATGGAGAGATCGCCAAC 140
DB 11 GlnIleGlnGlyValProPheProGluArgIleValLysTyrTrpThrArgValGluGln 30
QY 141 TTTCCGGTGGCGGCCCGAGCGTGTGGATCGTACCTACCCCAAGTCCGGCACCAGCTTG 200
DB 31 PheGlnAlaSerGluGluAspLeuIleAlaThrTyrProLysSerGlyThrThrTrp 50
QY 201 CTGAGGAGGTGGTCTACTTGGTGAGCCAGGGCGCTGACCCCGATGAGATCGGTTGATG 260
DB 51 IleGlnGluValValAspLeuIleLeuAsnAspGlyAspValAspLysCysLysArgAla 70
QY 261 AACATCGAGCAGCAGCTCCGCTCTGGAGTAC-----CCACAGCGGGC 305
DB 71 ProThrGlnValArgMetProPheLeuGluMetThrAlaAlaAspGlySerAsnAlaGly 90
QY 306 CTGACATCATCAAGAACTGACCTCTCCCGCTCATCAAGAGCACCTGCTCCCTACCGC 365
DB 91 IleThrLysLeuGluThrMetAspProProArgValIleLysThrHisLeuProIleGln 110
QY 366 TTTTGGCTCTGACCTCCAATGGAGACTCCCAAGGTCACTATATATGCTGCAACCCC 425
DB 111 LeuLeuProArgSerPheGluAsnAlaGlyCysLysValIleTyrValAlaArgAsnPro 130
QY 426 AAGGATCTGGTGGTCTTATTATCAGTTCCACCGCTCTCTCGGAGACCATGAGCTACCGA 485
DB 131 LysAspSerValValSerPheTyrHisPheAspArgMetAsnLeuArgGlnProGluPro 150
QY 486 GGCACCTTTCAAGAAATTCGCGGAGGTTTATGAATGATAAGCTGGCTACGGCTCCTGG 545
DB 151 GlyProTrpLysGlnTyrLeuGluArgPheMetLysGlyGlnLeuValTrpGlySerTrp 170
QY 546 TTTGAGCAGCTGACAGGTTCTGGGAGCACCAGCATGGACTCGAAGCTGCTTTTCTCAAG 605
DB 171 PheAspHisValLysGlyTyrTrpArgGluArgHisLysArgLysIleHisTyrMetPhe 190
QY 606 TATGAAGACATCATCGGACCTGGTGACGATGGTGAGCAGCTGGCCAGATCTCTGGGG 665
DB 191 PheGluAspMetLysGluAspProAlaArgGluValThrArgThrAlaGlnPheLeuGly 210
QY 666 GTGTCTGTGCAAGGCCAGCTGGAACCTCTGACGAGACCTGCGACCATGCTGGTG--- 722
DB 211 -----ArgGlnLeuSerGluSerThrIleAspHisIleValGlnMetThrThr 226
QY 723 -----GACCATGTGTGCAACCGCTGAGGCCCTGCC----- 752
DB 227 PheSerValMetArgGluAsnProMetAlaAsnTyrSerThrLeuProAspThrIlePhe 246
QY 753 -----GTGGCGCGGAGAGTGGGCTCTGGAAGACATCTTC 791
DB 247 AspArgThrAlaSerGlnPheMetArgLysGlyGluValGlyAspTrpLysAsnHisPhe 266
QY 792 ACCGCTCCATGAATGAGAAGTTTGACTGTGGTGTATAACAGAAAGTGGAAAGTGTGAC 851
DB 267 SerAlaGluGluAsnAlaAlaPheAspGluHisTyrGlnLysLysMetGluAspCysPro 286

Qy	855	ACGTTTACATTTTATTTATATACAGAACACACCACTCATCT	900
Db	285	LysPheillelrpAspLeu-AsnAspGluAlaAsnSerAsnHisSer	299
RESULT 97			
Q53SG4_HUMAN			
1D	Q53SG4_HUMAN	PRELIMINARY; PRT; 307 AA.	
AC	Q53SG4;		
DT	24-MAY-2005,	integrated into UniProtKB/TrEMBL.	
DT	24-MAY-2005,	sequence version 1.	
DT	07-FEB-2006,	entry version 3.	
DE	Hypothetical protein SUL1C1.		
GN	Name=SUL1C1;		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;		
OC	Homo.		
OX	NCBI_TaxID=9606;		
XP	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RP	Kalicki J., Drone K., Belter E.;		
RT	"The sequence of Homo sapiens BAC clone RP11-443K8.";		
RL	Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.		
RN	[2]		
RP	NUCLEOTIDE SEQUENCE.		
RA	Waterston R.H.;		
RL	Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.		
RN	[3]		
RP	NUCLEOTIDE SEQUENCE.		
RA	Waterston R.;		
RL	Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.		
RN	[4]		
RP	NUCLEOTIDE SEQUENCE.		
RA	Wilson R.K.;		
RL	Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.		
CC	-----		
CC	Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms		
CC	Distributed under the Creative Commons Attribution-NoDerivs License		
CC	-----		
DR	EMBL; AC019100; AAV14790.1; -; Genomic_DNA.		
DR	SMR; Q53SG4; 12-307.		
DR	Ensembl; ENSG00000198203; Homo sapiens.		
DR	GO; GO:0008146; F:sulfotransferase activity; IEA.		
DR	InterPro; IPR000863; Sulfotransferase.		
DR	Pfam; PF00685; Sulfotransferase 1; 1.		
DR	ProDom; PD001218; Sulfotransferase; 1.		
KW	Hypothetical protein.		
SO	SEQUENCE 307 AA; 35868 MW; DB83191FA88BFD1 CRC64;		
Alignment Scores:			
Pred. No.:	3,88e-19	Length:	307
Score:	395.50	Matches:	89
Percent Similarity:	49.2%	Conservative:	56
Best Local Similarity:	30.2%	Mismatches:	115
Query Match:	9.0%	Indels:	35
DB:	2	Gaps:	5
US-10-768-158-1 (1-2419) x Q53SG4_HUMAN (1-307)			
Qy	72	AAGPACTTCGAGTTCCATGCGGTCGCCGCCCTCTTCCCGCGGAAGATGGAGGAG	1
Db	12	LysLeuLysGluValGluGlyThrLeuLeuGlnProAlaThrValAspAsnTrpSerGln	3
Qy	132	ATGCCCAACTTCCCGTGGCGGCGAGCGTGTGGATCGTCACCTACCCCAAGTCGGC	1
Db	32	IleGlnSerPheGluAlaLysProAspAspLeuLeuIleCysThrTyPrProLysAlaGly	5
Qy	192	ACCAGCTTGCTGCAGGAGGTGGTCTACTTGGTGAGCCAGCGCGGTGACCCGATGATC	2
Db	52	ThrThrTrpIleGlnGluIleValAspMetIleGlnAsnGlyAspValGluLysCys	7
Qy	252	GGCTTGATGAACATCGACGAGCTCCCGGTCTCTGGAGTAC-----CCACAGCCG	3

```
Db 72 GlnArgAlaIleIleGlnHisArgHisProPheIleGluTrpAlaAargProProGlnPro 91
Qy 303 -----GGCCTGGACATCATCAAGAACTGACC 329
Db 92 SerGluThrGlyPheHisValAlaGlnAlaGlyLeuLysLeuLeuSerSerAen 111
Qy 330 TCTCCCGCCTCATCAAGAGCCACCTGCCCTACCGC---TTTCTGCCCTCTGACCTCCAC 386
Db 112 ProProAlaSerThrSerGlnSerAlaLysIleThrAspLeuLeuProProSerPheTrp 131
Qy 387 AATGAGACTCCAAGTCTATATGCTCGCAACCCCAAGGATCTGGTGTCTTAT 446
Db 132 GluAenAenCysAlaPheLeuTyrrValAlaAargAenAlaLysAAspCysMetValSerTyrr 151
Qy 447 TATCAGTTCCACCGCTCTCTGGGACCATGAGCTACCGAGGACCTTTCAGAAATTCGC 506
Db 152 TyrHisPheGlnAargMetAenHisMetLeuProAAspProGlyThrTrpGluGluTyrrPhe 171
Qy 507 CGGAGTTTATGAATGATAAGTACGGCTCTGCTGTGTGAGCAGCTGCAGGAGTTC 566
Db 172 GluThrPheIleAenGlyLysValValTrpGlySerTrpPheAspHisValLysGlyTrp 191
Qy 567 TGGGAGCACCGCATGAGTCAAGCTGCTTTTCTCAAGTATGAAGACATGATCGGCAC 626
Db 192 TrpGluMetLysAAspArgHisGlnIleLeuPheLeuPheTyrrGluAAspIleLysAargAAsp 211
Qy 627 CTGGTGACGATGGTGGAGCAGCTGCCAGATTCCTGGGGGTGCTCTGTGACAGGCCCCAG 686
Db 212 ProLysHisGluIleAargLysValMetGlnPheMetGlyLysLysValAAspGluThrVal 231
Qy 687 CTGGAAGCCCTGACGGAGCAC-----707
Db 232 LeuAAspLysIleValGlnGluThrSerPheGluLysMetLysGluAAsnProMetThrAAsn 251
Qy 708 -----TGCCACACGCTGTGGACCATGCTGCTCAAGCTGAGGCCCTGCCCCGTG 755
Db 252 ArgSerThrValSerLysSerIleLeuAAspGlnSerIleSerPhe-----Met 268
Qy 756 GCCCGGGAAGATTGGGTGTGGGAAGACATCTTCACCGTCTCCATGAATGAGAAGTTT 815
Db 269 ArgLysGlyThrValGlyAAspTrpLysAAsnHisPheThrValAlaGlnAAsnGluAargPhe 288
Qy 816 GACTTGGTGTATAACACAGAAGATGGAAAGTGTGACCTCACGTTT 860
Db 289 AspGluIleTyrrArgArgLysMetGluGlyThrSerIleAAsnPhe 303

RESULT 98
ID Q6DJ68_XENTR PRELIMINARY; PRT; 304 AA.
AC Q6DJ68_
DT 16-AUG-2004, integrated into UniProtKB/TrEMBL.
DT 16-AUG-2004, sequence version 1.
DE 07-FEB-2006, entry version 13.
DE MGC88979 protein.
GN Name=MGC88979;
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus; Silurana.
OX NCBI_TaxID=8364;
RN [1]_
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole body;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loguellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
```

```
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Patsy J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Trachman J., Schmutz J., Myers R.M.,
RA Butcherfield Y.S.N., Kraywinski M.I., Skalska U., Smallos D.E.,
RA Snerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole body;
RA Klein S., Gerhard D.S.; to the EMBL/GenBank/DBJ databases.
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; BC075315; AAH75315.1; -; mRNA.
DR Ensembl; ENSXETG0000009267; Xenopus tropicalis.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase 1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
SQ SEQUENCE 304 AA; 35586 MW; F73B991748AE2CFA CRC64;

Alignment Scores:
Pred. No.: 5,36e-19 Length: 304
Score: 393.50 Matches: 88
Percent Similarity: 54.8% Conservative: 56
Best Local Similarity: 33.5% Mismatches: 100
Query Match: 8.9% Indels: 19
DB: 2 Gaps: 5

US-10-768-158-1 (1-2419) x Q6DJ68_XENTR (1-304)
Qy 129 GAGATCGCAACTTCCCGGTGGCGGCCGAGCAGCGCTGGATCGTCACCTACCCCAAGTCC 188
Db 38 AspIleTrpAenPheArgAlaHisProGluAAspValLeuLeuAlaSerTyrrProLysAla 57
Qy 189 GGCACCAAGCTGTCTGACGAGGTGTCTACTTGTGTGAGCAGCGCGCTGACCCCGATGAG 248
Db 58 GlyIleThrTrpMetGlnGluValValAspMetIleTyrrGlnGluGlyAspThrAsnLys 77
Qy 249 ATCGGCTTGATGAACATCGACGACGACTCCCGTCTCTGGAGTAC-----CCACAGCG 302
Db 78 CysLeuAargAlaProThrTyrrAspArgHisPropheLeuGluAlaValProProLysPro 97
Qy 303 -----GGCCTGGACATCATCAAGAACTGACCTCTCCCGCCTCATCAAGAGCCAC 353
Db 98 ValProSerGlyLeuGlnLeuAlaGluMetGluProProAargValLeuLysThrHis 117
Qy 354 CTGCGCTACCGCTTCTCCCTCTGACCTCCACCAATGAGAGACTCCAAAGTCTATATATG 413
Db 118 LeuProIleGlnLeuIleProProSerPheTrpLysGlnAAsnCysLysValIleTyrrVal 137
Qy 414 GCTCGCAACCCCAAGATCTGGTGTGTCTTATATCATGTTCCACCGCTCTCTCGGAGCC 473
Db 138 AlaAargAenAlaLysAAspSerLeuValSerTyrrPheHisPheGlnAargMetThrLysGly 157
Qy 474 ATGAGTACCGAGGACCTTTCAAGAAATTCGCGGAGGTTTATGAATGAATAGCTGGGC 533
Db 158 LeuProAAspProGlyThrTrpAAspTyrrPheMetAlaPheLeuSerGlyThrLeuPro 177
Qy 534 TACGCTCTCTGTTGACACCGTTCAGGAGTCTGGGAGCACCGCATGGACTCGAACCTG 593
Db 178 TrpGlySerTrpPheAspHisValAAsnGlyTrpTrpLysAlaLysAAspHisArgVal 197
Qy 594 CTTTTCCTCAAGTATGAAGACATGATCGGACCTCGGACCTGTGTGAGCATGTGTGAGCAGCTGCC 653
Db 198 LeuTyrrValPheTyrrGluAAspMetLysAAspLeuAargLeuGluIleGlnAargVal 217
```

QY 654 AGATTCCTGGGGTGTCTGTGACAAAGCCAGCTGGAAGCCCTGACGAGCAC---TGC 710
 Db |||||
 218 SerPheLeuAspLysAspLeuProGluGluValLeuGluLysIleCysGlnHisThrThr 237
 QY 711 CACACAGCTGGTGACACG-----TGCTGCACAGCTGAGGCCCTGCCC----- 752
 Db |||||
 238 PheGlnAlaMetLysGluAsnProMetAlaAsnTyrThrThrMetProThrThrValMet 257
 QY 753 -----GTGGCGCGGGAAGATTGGGCTGTGGAAGACACATCTTC 791
 Db |||||
 258 AspGlnSerValSerProPheMetArgLysGlyIleValGlyAspTrpLysAsnHisPhe 277
 QY 792 ACCGCTCTCATGAATGAGAGTTTGACTGGTGTATAAACAGAGATGGAAGTGTGAC 851
 Db |||||
 278 LeuValAlaGlnAsnGluLeuPheAspTrpGluTyrLysArgMetAspGlyThrGly 297
 QY 852 CTCACGTTT 860
 Db |||||
 298 LeuAspPhe 300

RESULT 99
 ID 07ZYHO_XENLA PRELIMINARY; PRT; 304 AA.
 AC 07ZYHO;
 DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.
 DT 01-JUN-2003, sequence version 1.
 DT 07-FEB-2006, entry version 17.
 DE Sult1c1-prov protein.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;
 OC Xenopodinae; Xenopus; Xenopus.
 NC NCB1_TaxID=8355;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Embryo;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.F., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smalios D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.W., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Embryo;
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.;
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 RT initiative";
 RL Dev. Dyn. 225:384-391(2002).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Embryo;
 RX Klein S., Strausberg R.;
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.

CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC -----
 DR EMBL; BC043790; AAH43790.1; -, mRNA.
 DR HSSP; P50224; 1CJM.
 DR GO; GO:0008146; F:sulfotransferase activity; IEA.
 DR InterPro; IPR000863; Sulfotransferase.
 DR Pfam; PF00685; Sulfotransfer_1; 1.
 DR ProDom; PD001218; Sulfotransferase; 1.
 SQ SEQUENCE 304 AA; 35108 MW; 0D4C0C2B017C5692 CRC64;

Alignment Scores: 5.36e-19 Length: 304
 Pred. No.: 393.50 Matches: 95
 Score: 393.50
 Percent Similarity: 50.2% Conservative: 49
 Best Local Similarity: 33.1% Mismatches: 120
 Query Match: 8.9% Indels: 23
 DB: 2 Gaps: 5

US-10-768-158-1 (1-2419) x Q7ZYHO_XENLA (1-304)

QY 63 TTCGAGAGCAAGTACTTCGAGTTCATGGCTGGGGCTGCGGCCCTTCTCGCGCGGAAG 122
 Db |||||
 16 PheGlnValThrMetGlyHisIleGluGlyValProLeuProValThrThrCysAspMet 35
 QY 123 ATGGAGGAGATCGCAACTTCCCGTGGCGCCCGACGACGTGTGGATCGTCACTACCCC 182
 Db |||||
 36 TrpAspThrIleTyrAsnPheGlnAlaArgLysAspIleLeuIleAlaThrTyrPro 55
 QY 183 AAGTCCGCGCACCGTGTCTGTCAGAGGTGTCTACTTGTGTGAGCCAGGCGGTGAC--- 239
 Db |||||
 56 LysAlaGlyThrTrpMetGlnIleValAspLeuIleLeuGlnGluGlyAspVal 75
 QY 240 -----CCGATGAGATCGGCTTGATGAACATCGACGACGAGCTCCCG 281
 Db |||||
 76 GlnLysSerMetArgAlaProCysTyrIleLysValProPheIleAsp---LeuValPro 94
 QY 282 GTCCTGGAGTACCACAGCGCGGCTGACATCATCAAGAACTGACCTCTCCCGCCTC 341
 Db |||||
 95 ProLysProMetPro---ProGlyValAlaLeuAlaGlnThrMetAsnSerProArgIle 113
 QY 342 ATCAAGAGCCACTGCGCCCTACCGCTTCTGCGCTTCTGACCTCCACATGGAGACTCCAAG 401
 Db |||||
 114 LeuLysThrHisLeuProIleAsnLeuLeuProSerPheTrpGluLysAsnThrLys 133
 QY 402 GTCATCTATATGGCTCGCAACCCCAAGATCTGTGGTGTCTTATTATCAGTTCCACCGC 461
 Db |||||
 134 ValValTyrValAlaAlaArgAsnAlaLysAspSerMetValSerTyrTyrPheHisLys 153
 QY 462 TCTCTCGGACCATGAGCTACCGAGCACCTTTCAAGAAATTCGCGGAGGTTTATGAAT 521
 Db |||||
 154 MetAsnLysPheLeuProAspSerGlyThrLeuAspAsnPhePheSerGluPheLeuSer 173
 QY 522 GATAAGCTGGGTACGCTCTGTTGACACCTGACGAGTCTCTGGGACGACCGCATG 581
 Db |||||
 174 GlyAspValProTrpGlySerTrpPheAspAsnValLeuGlyTrpTrpLysAlaLeuAsp 193
 QY 582 GACTCGAAGCTGCTTTTCTCAAGATATCAAGACATCATCGGACCTGGTGTGACATGTG 641
 Db |||||
 194 LysHisGlnIleLeuPheIlePheTyrGluAspMetIleGlnAspProMetArgGluIle 213
 QY 642 GAGCAGCTGCCAGATTCTCTGGGGGTGTCTGTGACCAAGGCCAGCTGGAAGCCCTGACG 701
 Db |||||
 214 LysLysValMetThrPheLeuGlyLysAspLeuSerAspGluValLeuGluLysIleLys 233
 QY 702 GAGCAGCTGCCAC-----CAGCTGGTGACACGATGCTGCAACGCTGAGGCCCTGCC 752
 Db |||||
 234 TyrHisThrSerPheGlnAlaMetLysGluAsnProMetThrAsnAsnSerThrValPro 253
 QY 753 -----GTGGCGCGGGAAGAGTGTGGGCTGTGG 779
 Db |||||
 254 LysThrIleMetAspGlnThrIleSerProPheIleArgLysGlyThrValGlyAspTrp 273

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 17, 2006, 09:32:19 ; Search time 16.9 Seconds
(without alignments)
4131.628 Million cell updates/sec

Title: US-10-768-158-1
Perfect score: 4401
Sequence: 1 gcagcgcgacggcgccgc.....aaaaaaaaaaaaaaaaaaaaa 2419

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlp
-O=/abss/ABSSWEB spool/US10768158/runat 16052006 163424 5311/app query.fasta_1
-DB=PIR -OPWT=fastan -SUFFIX=n2p.rpr -MINMATCH=0.1 -LOOFC=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=100
-DOCALIGN=200 -THR SCORE=spt -THR MAX=100 -THR MIN=0 -ALIGN=100 -MODE=LOCAL
-OUTFMT=spt -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss02p
-USER=US10768158 @CGN 1.1 63 @runat 16052006 163424 5311 -NCPU=6 -ICPU=3
-NO_MMAP -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR 80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	495	11.2	298	2 S28183	aryl sulfotransfer
2	472	10.7	295	2 A55451	aryl sulfotransfer
3	467	10.6	296	2 JCT282	hydroxyarylamine s
4	458	10.4	291	2 S10329	aryl sulfotransfer
5	456	10.4	295	2 S52791	aryl sulfotransfer
6	456	10.4	295	2 JCS249	aryl sulfotransfer
7	456	10.4	295	2 G01843	aryl sulfotransfer
8	454	10.3	295	2 I57945	phenol-sulfating p
9	451	10.2	295	2 S52794	aryl sulfotransfer
10	449	10.2	295	2 JCS248	aryl sulfotransfer
11	449	10.2	295	2 JCS253	aryl sulfotransfer
12	447	10.2	296	2 JCT283	hydroxyarylamine s
13	446.5	10.1	299	2 JCS884	thyroid hormone su
14	446	10.1	295	2 G02924	phenol sulfotransf

15	446	10.1	295	2	S29045	estrone sulfotrans	
16	444	10.1	295	2	S52399	aryl sulfotransfer	
17	440.5	10.0	296	2	JCS885	thyroid hormone su	
18	435	9.9	295	2	I73679	estrogen sulfotran	
19	434	9.9	294	2	JC2229	estrogen sulfotran	
20	432.5	9.8	299	2	JEO186	sulfotransferase (
21	430	9.8	295	2	I53296	testis-specific es	
22	430	9.8	295	2	I56606	estrogen sulfotran	
23	429	9.7	295	2	A41930	estrone sulfotrans	
24	419.5	9.5	304	2	A49098	N-hydroxyarylamine	
25	413	9.4	295	2	JEO197	phenol sulfotransf	
26	412	9.4	294	2	JCS5000	aryl sulfotransfer	
27	408.5	9.3	296	2	A44011	adrenocortical est	
28	397.5	9.0	301	2	JW0078	amine sulfotransfe	
29	385.5	8.8	287	2	JCS4531	alcohol sulfotrans	
30	380	8.6	338	2	JEO196	hydroxysteroid sul	
31	376.5	8.6	287	2	A54026	alcohol sulfotrans	
32	371	8.4	286	2	JEO152	alcohol sulfotrans	
33	361.5	8.2	285	2	I55760	alcohol sulfotrans	
34	359.5	8.2	284	2	I60190	alcohol sulfotrans	
35	356.5	8.1	285	2	T10086	alcohol sulfotrans	
36	355.5	8.1	285	1	I38548	alcohol sulfotrans	
37	353.5	8.0	284	2	I52849	alcohol sulfotrans	
38	351.5	8.0	284	2	A35569	alcohol sulfotrans	
39	338	7.7	282	2	A26136	senescence marker	
40	303	6.9	311	2	B40216	flavonol 3'-sulfot	
41	302	6.9	329	2	T47448	sulfotransferase-1	
42	297.5	6.8	324	2	B84452	probable steroid s	
43	284.5	6.5	323	2	T47447	sulfotransferase-1	
44	280	6.4	325	2	T07833	probable steroid s	
45	279	6.3	326	2	A84452	probable steroid s	
46	278.5	6.3	346	2	E86319	probable flavonol	
47	276.5	6.3	323	2	T07831	probable steroid s	
48	275.5	6.3	320	2	A40216	flavonol 4'-sulfot	
49	267	6.1	350	2	H96768	protein flavonol s	
50	265.5	6.0	324	2	T07832	probable steroid s	
51	262.5	6.0	326	2	F86407	probable sulfotran	
52	261	5.9	302	2	S69188	probable flavonol	
53	254.5	5.8	296	2	JCT921	cytosolic sulfotra	
54	249.5	5.7	338	2	A96769	protein flavonol s	
55	240	5.5	333	2	A84523	probable steroid s	
56	235.5	5.4	314	2	T06012	hypothetical prote	
57	224.5	5.1	331	2	H84451	probable steroid s	
c	190.5	4.3	660	1	Q0BE3	BHLFI protein - hu	
59	161	3.7	280	2	A95936	probable alcohol s	
c	60	158	3.6	1446	1	A53344	immediate-early pr
c	61	157.5	3.6	1042	1	CGCH1S	collagen alpha 1(I
c	62	156.5	3.6	1460	1	EDBE1F	immediate-early pr
c	63	155.5	3.5	636	2	S41067	collagen alpha 1(I
c	64	154.5	3.5	392	1	PIHUB6	salivary proline-r
c	65	154.5	3.5	1958	2	B40505	hypothetical prote
c	66	153.5	3.5	358	1	WMBE38	infected cell prot
c	67	153.5	3.5	775	1	EDBE11	immediate-early pr
c	68	153.5	3.5	1453	2	S21626	collagen alpha 1(I
c	69	153	3.5	1414	1	S23809	collagen alpha 2(I
c	70	152	3.5	294	2	S3141	hypothetical prote
c	71	151.5	3.4	1464	2	S59856	collagen alpha 1(I
c	72	151.5	3.4	1791	2	T02345	hypothetical prote
c	73	151	3.4	2944	2	A54849	collagen alpha 1(I
c	74	150.5	3.4	1497	2	I49607	procollagen type V
c	75	150.5	3.4	3570	2	T45025	mucin MUC5B, trach
c	76	150	3.4	305	2	T20906	hypothetical prote
c	77	149.5	3.4	325	2	T32248	hypothetical prote
c	78	148.5	3.4	282	2	J50168	collagen col-8 - C
c	79	148.5	3.4	580	2	T43481	probable mucin DKF
c	80	148.5	3.4	784	2	JQ0317	hypothetical 82K p
c	81	148	3.4	3124	2	A40020	collagen alpha 1(X
c	82	147.5	3.4	2715	2	T13049	eyelid - fruit fly
c	83	147	3.3	310	1	PIHUSD	salivary proline-r
c	84	147	3.3	428	2	T24769	hypothetical prote
c	85	147	3.3	1603	2	S23810	collagen alpha 1(X
c	86	147	3.3	1691	1	S22917	collagen alpha 5(I
c	87	146.5	3.3	779	1	CGBO1S	collagen alpha 1(I


```

c 88      146      3.3      1039      2      T35878      hypothetical prote
c 89      145.5    3.3      251      1      PIHUPF      salivary proline-r
c 90      145      3.3      673      1      CGB06C      collagen alpha 1(I
c 91      145      3.3      1418      2      T45467      collagen alpha 1(I
c 92      144.5    3.3      283      2      T19732      hypothetical prote
c 93      144.5    3.3      286      2      S34665      collagen, cuticula
c 94      144.5    3.3      296      2      A31219      collagen 1 - Caeno
c 95      144      3.3      1487      1      CGHU6C      hypothetical alpha 1(I
c 96      143.5    3.3      301      2      T21314      hypothetical prote
c 97      143.5    3.3      316      2      T20497      hypothetical prote
c 98      143.5    3.3      333      2      T52355      hypothetical prote
c 99      143.5    3.3      671      1      CGR15      collagen alpha 1(I
100      143.5    3.3      1106      2      JQ0405      hypothetical 119.5

RESULT 1
S28183
A:Title: aryl sulfotransferase (EC 2.8.2.1) p1 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C:Accession: S28183
R:Kong, A.N.T.; Ma, M.; Tao, D.; Yang, L.,
Biochim. Biophys. Acta 1171, 315-318, 1993
A:Title: Molecular cloning of cDNA encoding the phenol/aryl form of sulfotransferase (mg
A:Reference number: S28183, MUID:93144346; PMID:8424956
A:Accession: S28183
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-298 <KON>
A:Cross-references: UNIPROT:Q9R155; UNIPARC:UPI0000175825
A:Note: the sequence from Fig. 2 is inconsistent with that from Fig. 1 in having 51-Tyr
C:Superfamily: alcohol sulfotransferase
C:Keywords: sulfotransferase

Alignment Scores:
Pred. No.:      Length:      298
Score:          495.00      Matches:      104
Percent Similarity: 53.9%      Conservative: 54
Best Local Similarity: 35.5%      Mismatches:  117
Query Match:    11.2%      Indels:       18
DB:              Gaps:       3

US-10-768-158-1 (1-2419) x S28183 (1-298)

QY      36      GCGGAGACCCCGACGACCCCGGGGAGTTCGAGAGCAAGTACTTCGAGTTCATGCGCGT 95
Db      2      AlaGlnAsnProSerAsnMetGluProLeuArgLysProLeuValProVallylsGlylie 21
QY      96      CGGCTGCCGCCCTTCTTCGCCGGGAGATGCGAGGAGATGCCCAACTTCCCGTGGCGGCC 155
Db      22      ProLeuileLysTyrPheAlaGluThrMetGluGlnLeuGlnAsnPheThrAlaTyrPro 41
QY      156      ACGGAGCTGTGATCTGCTACCTACCCCAAGTCGGGACGAGCTTGTCTGAGGAGGTGTC 215
Db      42      AspAspValLeuileSerThrTyrProLysSerGlyThrTrpMetSerGluileMet 61
QY      216      TACTTGGTGACGCGGCGCTGACCCCGATGAGATCGGCTTGATGAACATCGACGAGCAG 275
Db      62      AspMetileTyrGlnGlyLysLeuAspLysCysGlyArgAlaProValTyrAlaArg 81
QY      276      CTCGCCGCTCTTGAGTACCCCA-----CAGCCGGGCTGGACATCATCAAGGAA 323
Db      82      IleProPheLeuGluPheSerCysProGlyValProProGlyLeuGluThrLeuLysGlu 101
QY      324      CTGACCTCTCCCGCTCATCAAGAGCCAGCTCCCTACCGCTTTCGCGCTCTGACCTC 383
Db      102      ThrProAlaProArgileileLysThrHisLeuProLeuSerLeuLeuProGlnSerLeu 121
QY      384      CACAATGGAGACTCCCAAGGTCTATATATGCTCGCAACCCCAAGGATCTGGTGGTGTCT 443
Db      122      LeuAspGlnLysileLysValIleTyrValAlaArgAsnAlaLysAspValValSer 141

```

```

QY      444      TATTATCAGTTCCACCGCTCTCTCGGGACCATGAGTACCGAGGACACCTTTCAAGAATTC 503
Db      142      TyrTyrAsnPheTyrLysMetAlaLysLeuHisProAspProGlyIleThrTrpGluSerPhe 161
QY      504      TGCCGGAGTTTATGAATGATAAGCTGGGTACGCTCCTGTTGTGACGACGCTGCAGGAG 563
Db      162      LeuGluAsnPheMetAspGlyLysValSerTyrGlySerTrpTyrGlnHisValysGlu 181
QY      564      TTCTGGGAGCACCGCATGAGACTCGAACGTGCTTTTCTCAAGTATGAAGACATGCATCGG 623
Db      182      TrpTrpGluLeuArgArgThrHisProValLeuTyrLeuPheTyrGluAspMetLysGlu 201
QY      624      GACCTGGTGACATGTGTGACGACGCTGCCAGATTCCTGGGGGTGCTCTGTGACAAGGCC 683
Db      202      AsnProLysArgGluileLysLysIleLeuGluPheLeuGlyArgSerLeuProGluGlu 221
QY      684      CAGCTGGAAGCCCTGACGAGCACATGC-----CACCAGCTGGTGGACACGATGCTGC 734
Db      222      ThrValAspLeuileValHisThrSerPheLysLysMetLysGluAsnProMetAla 241
QY      735      AACGCTGAGGCCCTGCCCC-----GTGGGCGCG 761
Db      242      AsnTyrThrThrIleProThrGluValMetAspHisThrIleTyrProPheMetArgLys 261
QY      762      GGAAGAGTTGGCTGTGAAGACATCTTCACCGCTCCATGATGAAGAGTTTGCATCTG 821
Db      262      GlyThrIleGlyAspTrpLysAsnThrPheThrValAlaGlnSerGluHisPheAspAla 281
QY      822      GTGTATTAACAGAGATCGGAAAGTGTGACCTCACGTTT 860
Db      282      HisTyrAlaLysLeuMetThrGlyCysAspPheThrPhe 294

RESULT 2
A55451
A:Title: aryl sulfotransferase (EC 2.8.2.1) HAST3 / estrogen sulfotransferase EST - human
N:Alternate names: gene STM protein; thermolabile phenol sulfotransferase
C:Species: Homo sapiens (man)
C:Date: 10-Feb-1995 #sequence_revision 10-Feb-1995 #text_change 09-Jul-2004
C:Accession: A55451; JN0713; JC2067; S52398; I38551; I57679; I52299; PC2031
J:Bernier, F.; Leblanc, G.; Labrie, F.; Luu-The, V.
J. Biol. Chem. 269, 28200-28205, 1994
A:Title: Structure of human estrogen and aryl sulfotransferase gene. Two mRNA species i-
A:Reference number: A55451; MUID:95050600; PMID:7961757
A:Accession: A55451
A:Molecule type: DNA
A:Residues: 1-295 <BER>
A:Cross-references: UNIPROT:PS0224; UNIPARC:UPI0000046BED; GB:I34160
R:Zhu, X.; Veronese, M.E.; Bernard, C.C.A.; Sansom, L.N.; McManus, M.E.
Biochem. Biophys. Res. Commun. 195, 120-127, 1993
A:Title: Identification of two human brain aryl sulfotransferase cDNAs.
A:Reference number: JN0713; MUID:93371391; PMID:8363592
A:Accession: JN0713
A:Molecule type: mRNA
A:Residues: 1-295 <ZHU>
A:Cross-references: UNIPARC:UPI0000046BED; GB:L19956; NID:9306456; PIDN:AAA02943.1; PID
R:Wood, T.C.; Aksoy, I.A.; Aksoy, S.; Weinshilboum, R.M.
Biochem. Biophys. Res. Commun. 198, 1119-1127, 1994
A:Title: Human liver thermolabile phenol sulfotransferase: cDNA cloning, expression and
A:Reference number: JC2067; MUID:94161723; PMID:8117269
A:Accession: JC2067
A:Molecule type: mRNA
A:Residues: 1-295 <WO>
A:Cross-references: UNIPARC:UPI0000046BED; GB:U08032; NID:9468256; PIDN:AAA17723.1; PID
R:Jones, A.L.; Hagen, M.; Coughtrie, M.W.H.; Roberts, R.C.; Glatt, H.R.
submitted to the EMBL Data Library, February 1995
A:Description: Human platelet phenol/sulfotransferases: cDNA cloning, stable expression
A:Reference number: S52398
A:Accession: S52398
A:Status: preliminary
A:Molecule type: mRNA

```


A:Residues: 1-295 <JON>
A:Cross-References: UNIPARC:UPI0000046BED; EMBL:X84653; NID:G671533; PIDN:CAA59146.1; PI
R:Aksoy, I.A.; Callen, D.F.; Apostolou, S.; Her, C.; Weinshilboum, R.M.
Genomics 23, 275-277, 1994
A:Title: Thermolabile phenol sulfotransferase gene (STM): Localization to human chromosome
A:Reference number: AS5725; MUID:95130098; PMID:7829089
A:Accession: I38551
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 139-198 <RES>
A:Cross-References: UNIPARC:UPI000016A097; EMBL:U08099; NID:G488284; PIDN:AAA82126.1; PI
R:Bernier, F.; Lopez-Solache, I.; Labrie, F.; Luu-The, V.
Mol. Cell. Endocrinol. 99, 11-15, 1994
A:Title: Cloning and expression of cDNA encoding human placental estrogen sulfotransferase
A:Reference number: I57679; MUID:94244847; PMID:8187952
A:Accession: I57679
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-295 <RE2>
A:Cross-References: UNIPARC:UPI0000046BED; GB:I25275; NID:G463124; PIDN:AAA36523.1; PID:
R:Dooley, T.P.; Probst, P.; Munroe, P.B.; Mole, S.E.; Liu, Z.; Doggett, N.A.
Biochem. Biophys. Res. Commun. 205, 1325-1332, 1994
A:Title: Genomic organization and DNA sequence of the human catecholamine-sulfating phen
A:Reference number: I52299; MUID:95100963; PMID:7802665
A:Accession: I52299
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-295 <RE3>
A:Cross-References: UNIPARC:UPI0000046BED; GB:S74971; NID:G833898
C:Comment: This enzyme transfers a sulfonate moiety from 3'-phosphoadenosine-5-phosphosul
C:Comment: Brain aryl sulfotransferase and placental estrogen sulfotransferase are trans
C:Comment: differ only in their 5'-untranslated regions.
C:Genetics:
A:Gene: GDB:STE
A:Cross-References: GDB:342104; OMIM:600043
A:Map position: 16p11.2-16p11.2
A:Introns: 50/1; 92/1; 124/3; 167/1; 198/3; 259/1
C:Superfamily: alcohol sulfotransferase
C:Keywords: brain; placenta; sulfotransferase

Alignment Scores:
Pred. No.: 1,09e-27 Length: 295
Score: 472.00 Matches: 100
Percent Similarity: 54.5% Conservative: 52
Best Local Similarity: 35.8% Mismatches: 101
Query Match: 10.7% Indels: 26
DB: 2 Gaps: 4

US-10-768-158-1 (1-2419) x A55451 (1-295)

QY 90 GCGGTGCGGCTGCGGCTTCTGCGCGGGAAGATGGAGAGATCGCCAACTTCCGGTGT 149
DB 17 GlyValProLeuIleLysTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla 36

QY 150 CGGCCAGGACGTGGTGGATGTCACCTACCCCAAGTCGGACACGCTTGTCTGCGAGAG 209
DB 37 ArgProAspAspLeuIleAsnThrTyrProLysSerGlyThrThrTyrPheValSerGln 56

QY 210 GTGGTCTACTTGGTGGACGCGGCTGACCCCGATGAGATCGGCTGTGATGAACATCGAC 269
DB 57 IleLeuAspMetIleTyrGlnGlyAspLeuGlyLysCysAsnArgAlaProIleTyr 76

QY 270 GAGCAGCTCCGGCTCTGGAGTACCCACAGCGG-----GGCCTGGACATCATC 317
DB 77 ValArgValProPheLeuGluValAsnAspProGlyGluProSerGlyLeuGluThrLeu 96

QY 318 AAGGAAGTACCTCTCCCGCTCATCAAGAGCCACTGCGCTTGTGCGCTTCTGCGCTCT 377
DB 97 LysAspThrProProArgLeuIleLysSerHisLeuProLeuAlaLeuLeuProGln 116

QY 378 GACCTCCAAATGGAGATCCCAAGTCACTATATGGCTCGAACCCCAAGCATCTGGTG 437
DB 117 ThrLeuLeuAspGlnLysValLysValValTyrValAlaArgAsnProLysAspValAla 136

QY 438 GTGTCTTATTATCACTTCAACCGCTCTCTGCGGACCATGAGCTACCGAGGACACTTTTCAA 497
DB 137 ValSerTyrTyrHisPheHisArgMetGluLysAlaHisProGluProGlyThrTyrPhe 156
QY 498 GAATTTCTCCGAGGAGTTTATCAATGATGAAGCTGGGCTCGGCTCTGTTTGGAGCACGTG 557
DB 157 SerPheLeuGluLysPheMetAlaGlyGluValSerTyrGlySerTyrPheGlnHisVal 176
QY 558 CAGGAGTTCTGGGAGCACCGCATGCACTCGAACGTGCTTTTTTCTCAAGTATGAAGACATG 617
DB 177 GlnGluTyrTyrGluLeuSerArgThrHisProValLeuTyrLeuPheTyrGluAspMet 196
QY 618 CATCGGACCTGCTGATGACCATGCTGGAGCAGCTGGCCAGATTCCTGGGGGTGCTCTGTGAC 677
DB 197 LysGluAsnProLysArgGluIleGlnLysIleLeuGluPheValGlyArgSerLeuPro 216
QY 678 AAGGCCAGCTGGAAGCCCTGACGAGCAGCTGC----- 710
DB 217 GluGluThrMetAspPheMetValGlnHisThrSerPheLysGluMetLysLysAsnPro 236
QY 711 -----CACCAGCTGGTGACCACTGTGTGCAACCTGAGGCC 746
DB 237 MetThrAsnTyrThrThrValProGlnGluLeuMetAspHisSerIleSer----- 253
QY 747 CTGCCC-----GTGGCGCGGGAAGCTTGGGCTGTGGAAGACATCTTACCGCTCTCCATG 803
DB 254 ---ProPheMetArgLysGlyMetAlaGlyAspTyrLysThrPheThrValAlaGln 272
QY 804 AATGAGAAGTTGACTTGGTGTATAAACAGAGATGGGAAAGTGGACCTCAGCTTT 860
DB 273 AsnGluArgPheAspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe 291

RESULT 3
JC7282
hydroxyarylamine sulfotransferase (EC 2.8.2.-) 2 - rat
N:Alternate names: SULT1C2 sulfotransferase
C:Species: Rattus norvegicus (Norway rat)
C:Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 09-Jul-2004
C:Accession: JC7282
R:Li, X.; Joehnk, C.; Hartmann, D.; Schestag, F.; Kroemer, W.; Gieselmann, V.
Biochem. Biophys. Res. Commun. 272, 242-250, 2000
A:Title: Enzymatic properties, tissue-specific expression, and lysosomal location of tw
A:Reference number: JC7282
A:Contents: Kidney
A:Accession: JC7282
A:Molecule type: mRNA
A:Residues: 1-296 <LIA>
A:Cross-References: UNIPROT:Q9WUW8; UNIPARC:UPI00001360FC; GB:AJ238391
C:Genetics:
A:Gene: sult1c2
C:Superfamily: alcohol sulfotransferase
C:Keywords: kidney; sulfotransferase

Alignment Scores:
Pred. No.: 2.6e-27 Length: 296
Score: 467.00 Matches: 100
Percent Similarity: 55.4% Conservative: 58
Best Local Similarity: 35.1% Mismatches: 105
Query Match: 10.6% Indels: 22
DB: 2 Gaps: 5

US-10-768-158-1 (1-2419) x JC7282 (1-296)

QY 66 GAGACCAAGTACTTCCAGTTCATGCTGCGGCTGCCGCCCTTCTGCGCGGGAAGATG 125
DB 10 GlnThrLysLeuLysGluValAlaGlyIleProLeuGlnAlaProThrValAspAsnTrp 29

QY 126 GAGGAGATCGCAACTTCCCGTGGCGGCCACGACGTGTGGATCGTCACTACCCCAAG 185
DB 30 SerGlnIleGlnThrPheLysAlaLysProAspLeuLeuIleCysThrTyrProLys 49

QY 186 TCCGACACAGCTTCTGTCAGGAGGTGTCTACTTGGTGAGCCAGGCGGCTGACCCCGAT 245

RESULT 5

S52791
aryl sulfotransferase (EC 2.8.2.1) - human
C:Species: Homo sapiens [man]
C>Date: 19-May-1995 #sequence_revision 21-Jul-1995 #text_change 18-Jun-1999
C:Accession: S52791
R:Ozawa, S.; Negata, K.; Shimada, M.; Ueda, M.; Tsuzuki, T.; Yamazoe, Y.; Kato, R.
submitted to the EMBL Data Library, March 1994
A:Description: Primary structures and properties of two related aryl sulfotransferases
A:Reference number: S52791
A:Accession: S52791
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-295 <OZA>
A:Cross-references: UNIPARC:UPI000016B06E; EMBL:X78282; NID:g758594; PIDN:CAA55088.1; P:
C:Superfamily: alcohol sulfotransferase
C:Keywords: sulfotransferase

Pred. NO.:	1.78e-26	Length:	295
Score:	456.00	Matches:	94
Percent Similarity:	54.5%	Conservative:	58
Best Local Similarity:	33.7%	Mismatches:	101
Query Match:	10.4%	Indels:	26
DB:	2	Gaps:	4

US-10-768-158-1 (1-2419) x S52791 (1-295)

QY	90	GGCGTGGGCTGC	CCCTTCGCGCGGGAAGATGGAGGATCGCAACTCCCGGTG	149
Db	17	GlyValLeuLeuIleLysTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnala	36	
QY	150	CGGCCACGACGCTGTGCTGCCTCACCTACCCTCCGAGCCGACCATGCTGCTCGACGAG	209	
Db	37	ArgProAspLeuLeuIleSerThrTyProllysSerGlyThrTrpValSerGln	56	
QY	210	GTGCTCTACTTTGTGAGCCAGGGCGCTGACCCCAGATGAGATCGGCTTGATGAACATCGAC	269	
Db	57	IleLeuAspMetIleTyrGlnGlyGlyAspLeuGluLysCysHisArgAlaProIlePhe	76	
QY	270	GAGCAGCTCCGGTCTGGAGTAGTACCCACACGCG-----GGCCTGGACATCATC	317	
Db	77	MetArgValProPheLeuGluPheLysValProGlyIleProSerGlyMetGluThrLeu	96	
QY	318	AAGGAACCTGACCTCTCCCGGCTCATCAAGAGCCACCTGCCTACCGCTTCTGCGCCTCT	377	
Db	97	LysAsnThrProAlaProArgLeuLeuIleThrsHisLeuProLeuAlaLeuLeuProGln	116	
QY	378	GACCTCCACAATGAGACTCCAAGTTCATCTATATGCTCGCAACCCCAAGGATCTGGT	437	
Db	117	ThrLeuLeuAspGlnLysValLysValValTyrValAlaArgAsnAlaLysAspValala	136	
QY	438	GTGCTCTATTATCATGTTCCACCGCTCTCTCGCGACCATGAGCTACCGAGGACCTTTCAA	497	
Db	137	ValSerTyrTyrHisPheTyHisMetAlaLysValTyrProHisProGlyThrTrpGlu	156	
QY	498	GAATTCCTCCGAGGTTTATGAATGATAAGTACGGCTACGGCTCCTCGTTTGAGCAGGTG	557	
Db	157	SerPheLeuGluLysPheMetAlaGlyGluValSerTyrGlySerTyrTyrGlnHisVal	176	
QY	558	CAGAGTTCTGGGAGCACCCGATCGGACTCGAACGCTGTTTTTCTCAAGTATGAAGACATG	617	
Db	177	GlnGluTrpTrpGluLeuSerArgThrHisProValLeuTyrLeuPheTyrGluAspMet	196	
QY	618	CATCGGACCTGGTGCAGATGGTGGAGCAGCTGGCCAGATCTCTGGGGGTGCTCTGTGAC	677	
Db	197	LysGluAsnProTyrLysArgGluIleGlnLysIleLeuGluPheValGlyArgSerLeuPro	216	
QY	678	AAGGCCACGCTGGAAGCCCTGACGAGCAGTGC-----CACCAGCTGGTGACCAAGTCTGCAACGCTGAGGCC	746	
Db	217	GluGluThrValAspLeuMetValGluHisThrSerPheLysGluMetLysLysAsnPro	236	
QY	711	-----CACCAGCTGGTGACCAAGTCTGCAACGCTGAGGCC	746	

```

Db      237 MetThrAsnTyrThrThrValArgGluPheMetAspHisSerIleSer----- 253
      747 CTGCCC---GTGGGCGGGGAGAGCTTGGCGTGTGGAAGACATCTTACCGTCTCCATG 803
Db      254 ---ProPheMetArgGlyGlyMetAlaGlyAspTrpLysThrThrPheThrValAlaGln 272
Qy      804 AATGAGAAATTTGACTTGGTGTATAAACAGAGAAGATGGGAAGCTGACCTCAGTTT 860
Db      273 AsnGluArgPheAspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe 291

RESULT 6
JC5249
N:Alcarnate names: phenol sulfotransferase; sulfokinase
C:Species: Homo sapiens (man)
C:Date: 13-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: JC5249
R:Dooley, T.P.; Huang, Z.
Biochem. Biophys. Res. Commun. 228, 134-140, 1996
A:Title: Genomic organization and DNA sequences of two human phenol sulfotrans-
A:Reference number: JC5248; MUID: 97069665; PMID: 8912648
A:Accession: JC5249
A:Molecule type: mRNA
A:Residues: 1-295 <DOO>
A:Cross-references: UNIPROT:P50226; UNIPARC:UPI0000136229; GB:U76619; NID: g1566
C:Comment: This enzyme catalyzes the transfer of the sulfonate group from phos-
C:Genetics:
A:Gene: stp2
A:Map position: 16p12.1-pl1.2
C:Superfamily: alcohol sulfotransferase
C:Keywords: sulfotransferase

Alignment Scores:
Pred. No.: 1,786-26 Length: 295
Score: 456.00 Matches: 94
Percent Similarity: 54.5% Conservative: 58
Best Local Similarity: 33.7% Mismatches: 101
Query Match: 10.4% Indels: 26
DB: 2 Gaps: 4

US-10-768-158-1 (1-2419) x JC5249 (1-295)
Qy      90 GCGCTGCGCTCGCCCTCTTCCGCGGGAAGATGAGGAGATGCCCAACTTCCCGTG 149
Db      17 GlyValProLeuIleLeTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla 36
Qy      150 CGGCCCCAGCAGCTGTGGATCGTCACTACCCCAAGTCCGGCACCAGCTGTGTGACAGGAG 209
Db      37 ArgProAspAspLeuLeuIleSerThrTyrProLysSerGlyThrThrTrpValSerGln 56
Qy      210 GTGCTCTACTTGGTGAGCCAGGGCGCTCACCCGATGAGATCGGCTTGATGAACATCGAC 269
Db      57 IleLeuAspMetIleTyrGlnGlyGlyAspLeuGluLysCysHisArgAlaProIlePhe 76
Qy      270 GAGCAGCTCCGCTCTCGAGTACCCACAGCCG-----GGCCTGGACATCATC 317
Db      77 MetArgValProPheLeuGluPheLysValProGlyIleProSerGlyMetGlnThrLeu 96
Qy      318 AAGGAACCTGACCTCTCCCGGCTCATCAAGACCCACTGCCTACCGCTTCTGCGCCTCT 377
Db      97 LysAsnThrProAlaProArgLeuLeuLysThrHisLeuProLeuAlaLeuLeuProGln 116
Qy      378 GACCTCCCAATGGAGACTCAAGGTCTATATATGGCTGCCAACCCCAAGGATCTGGTG 437
Db      117 ThrLeuLeuAspGlnLysValLysValValTyrValAlaIa-rGAsnAlaLysaspValAla 136
Qy      438 GTGTCTTATTATCAGTCTCCACCGCTCTCTCGCGGACCATGAGCTACCGAGGCACTTTCAA 497
Db      137 ValSerTyrThrHisPheTyrHisMetAlaLysValTyrProHisProGlyThrTrpGlu 156
Qy      498 GAATTCTCGCGAGGTTTATGAATGATTAAGCTACGGCTACGGCTCTCTGTTTGTGACACGTG 557

```

```

Db      157 SerPheLeuGluLysPheMetAlaGlyGluValSerTyrGlySerTrpTyrGlnHisVal 176
Qy      558 CAGGAGTTCTGGGACGACCAGCATGTGCAGACGTGGCTTTTTTCTCAAGTATGAAGACATG 617
Db      177 GlnGluTrpTrpGluLeuSerArgThrHisProValLeuTyrLeuPheTyrGluAspMet 196
Qy      618 CATCGGAGCCTGGTGACGATGGTGTGAGCAGTCGCCAGATTCTCTGGGGTGTCCCTGTGAC 677
Db      197 LysGluAenProLysArgGluIleGlnLysIleLeuGluPheValGlyArgSerLeuPro 216
Qy      678 AAGGCCACAGCTGGAAGCCCTGACGCGACTGC----- 710
Db      217 GluGluThrValAspLeuMetValGluHisThrSerPheLysGluMetLysLysThrPro 236
Qy      711 -----CACACAGCTGGTGGACCACTGTGTGCAACGCTCAGGCC 746
Db      237 MetThrAsnTyrThrThrValArgGluPheMetAspHisIleSer----- 253
Qy      747 CTGCCC---GTGGGGCCGGGAAGATTGGCTGTGTGAAGGACATCTTACCTCTCCCATG 803
Db      254 --ProPheMetArgLysGlyMetAlaGlyAspTrpLysThrPheThrValAlaGln 272
Qy      804 AATCAGAAGATTGACTGGTGTATAAACAGAGATGGGAAAGTGTGCACCTCACGTTT 860
Db      273 AsnGluArgPheAspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe 291

RESULT 7
G01843
aryl sulfotransferase - human
C/Species: Homo sapiens (man)
C/Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 09-Jul-2004
C/Accession: G01843
R/Zhu, X.
submitted to the EMBL Data Library, May 1995

```

QY	270	GAGCAGCTCCCGTCTCTGGAGTACCCACAGCCG-----GGCTGGACATCATC	317
Db	77	MetArgValProPheLeuGluPheLysAlaProGlyIleProSerGlyMetGluThrLeu	96
QY	318	AAGGAACGACCTCCCGCTCATCAAGAGCCACCTGCCCTACCGCTTTCTGCCCTCT	377
Db	97	LysAspThrProAlaProArgLeuLeuLysThrHisLeuProLeuAlaLeuProGln	116
QY	378	GACCTCCCAATGGAGACTCCAAGTCACTATATATGGCTGCGCAACCCCAAGATCTGTG	437
Db	117	ThrLeuLeuAspGlnLysValLysValValTyrValAlaArgAsnAlaLysAspValAla	136
QY	438	GTGCTCTATTATCAGTTTCCACCGCTCTCTCGGACCATTGAGCTACCGAGGACCTTTCAA	497
Db	137	ValSerTyrTyrHisPheTyrHisMetAlaLysValHisProGluProGlyThrTrpAsp	156
QY	498	GAATTCTCCCGAGGTTTATGAATGATAAGCTGGGCTACGGCTCTCTGTTTGGACACGTG	557
Db	157	SerPheLeuGluLysPheMetValGlyGluValSerTyrGlySerTyrGlnHisVal	176
QY	558	CAGGAGTTCTGGGACACCGCATGGACTCGAACGTGCTTTTTCACAGTATGAAGACATG	617
Db	177	GlnGluTrpTrpGluLeuSerArgThrHisProValLeuTyrLeuPheTyrGluAspMet	196
QY	618	CATCGGACCTGGTCACCATGGTGGAGCAGCTGGCCAGATCTCTGGGGGTCTCTGTGCAC	677
Db	197	LysGluAenProLysArgGluIleGlnLysIleLeuGluPheValGlyArgSerLeuPro	216
QY	678	AAGGCCACGCTGGAGGCCCTGACGGAGCACTGC-----CACCAGCTGGTGGACCAG	728
Db	217	GluGluThrValAspPheMetValGlnHisThrSerPheLysGluMetLysLysAsnPro	236
QY	729	TGCTGCAACGCTGAGGCCCTCGCC-----GTG 755	
Db	237	MetThrAenTyrThrThrValProGlnGluPheMetAspHisSerIleSerProPheMet	256
QY	756	GGCCGGGGAAGAGTTGGCTCTGGAAGGACATCTTCACCGTCTCCATCAATGAGAAGTTT	815
Db	257	ArgLysGlyMetAlaGlyAspTrpLysThrThrPheThrValAlaGlnAsnGluArgPhe	276
QY	816	GACTTGTGTATAAACAAGAGTGGAAAGTGTGCACCTCACGTTTT 860	
Db	277	AspAlaAspTyrAlaLysLysMetAlaGlyCysSerLeuThrPhe 291	
RESULT 9			
S52794			
aryl sulfotransferase (EC 2.8.2.1) - human			
C/Species: Homo sapiens (man)			
C/Date: 19-May-1995 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004			
C/Accession: S52794			
R/Ozawa, S.; Nagata, K.; Shimada, M.; Ueda, M.; Tsuzuki, T.; Yamazoe, Y.; Kato, R.			
submitted to the EMBL Data Library, March 1994			
A/Description: Primary structures and properties of two related aryl sulfotransferases			
A/Reference number: S52791			
A/Accession: S52794			
A/Status: preliminary			
A/Molecule type: mRNA			
A/Residues: 1-295 <OZA>			
A/Cross-references: UNIPROT:P50225; UNIPARC:UPI0000136228; EMBL:X78283; NID:g758596; PII:			
C/Superfamily: alcohol sulfotransferase			
C/Keywords: sulfotransferase			
Alignment Scores:			
Pred. No.:	4.27e-26	Length:	295
Score:	451.00	Matches:	93
Percent Similarity:	54.5%	Conservative:	57
Best Local Similarity:	33.8%	Mismatches:	107
Query Match:	10.2%	Indels:	18
DB:	2	Gaps:	3
US-10-768-158-1 (1-2419) x S52794 (1-295)			

QY 90 GCGGRCGGCTGCCGCCCTTCTGCGCGGGAGATGGAGGAGATGCCAACTTCCCGGTG 149

Db 17 GlyValProLeuIleLysTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla 36

QY 150 CGGCCACGCGCTGTGATCGTCACCTACCCCAAGTCCGGCACCAGCTTGTGCAGGAG 209

Db 37 ArgProAspAspLeuLeuIleSerThrTyrProLysSerGlyThrThrTrpValSerGln 56

QY 210 GTGGTCTACTTGGTGAGCAGCGCGCTGACCCCGATGAGATCGGCTTGTGATGAACATCGAC 269

Db 57 IleLeuAspMetIleTyrGlnGlyGlyAspLeuGluLysCysHisArgAlaProIlePhe 76

QY 270 GAGCAGCTCCGGTCTCGGATGATCCACCGCG-------GGCTGGACATCATC 317

Db 77 MetArgValProPheLeuGluPheLysAlaProGlyIleProSerGlyMetGluThrLeu 96

QY 318 AAGGAACATGACCTCTCCCGGCTCATCAAGACCCACCTGCCCTACCGCTTTCGGCCCTCT 377

Db 97 LysAspThrProAlaProArgLeuLeuLysThrHisLeuProLeuAlaLeuLeuProGln 116

QY 378 GACCTCCAAATGGAGACTCCAAAGTCACTATATGGCTCGCAACCCCAAGGATCTGGTG 437

Db 117 ThrLeuLeuAspGlnLysValLysValValTyrValAlaAlaArgAsnAlaLysAspValAla 136

QY 438 GTGCTCTAATATCATGTTCCACCGCTCTCTCGGGACCATGAGTACGAGGCGACCTTTCAA 497

Db 137 ValSerTyrTyrHisPheTyrHisMetAlaLysValHisProGluProGlyThrTrpAsp 156

QY 498 GAATTCCTCCGGAGGTTTATGAATGATAAAGCTGGGCTACGGCTCTCTGGTTTTCAGACATG 557

Db 157 SerPheLeuGluLysPheMetValGlyGluValSerTyrGlySerTyrPheGlnHisVal 176

QY 558 CAGGATGTTCTGGGAGCACCGCATGGACTCGAACGTCGCTTTTCTCAAGTATGAAGACATG 617

Db 177 GlnGluTrpTrpGluLeuSerArgThrHisProValLeuTyrLeuPheTyrGluAspMet 196

QY 618 CATCGGACCTGGTGACGATGGTGAGCAGCTGGCCAGATTCCTGGGGGTGCTGTGTGAC 677

Db 197 LysGluAsnProLysArgGluIleGlnLysIleLeuGluPheValGlyArgSerLeuPro 216

QY 678 AAGGCCACGCTGGAAGCCCTCAGCGGACGACTGC-----CACCAGCTGGTGGACCAG 728

Db 217 GluGluThrValAspPheMetValGlnHisThrSerPheLysGluMetLysLysAsnPro 236

QY 729 TGCTGCAACGCTGAGGCCCTCGCC-----GTG 755

Db 237 MetThrAsnTyrThrValProGlnPheMetAspHisSerIleSerProPheMet 256

QY 756 GCGCGGGAAGATTGGGCTGTGGAAGACATCTTCACCGTCTCCATGAATGAGAAGTTT 815

Db 257 ArgLysGlyMetAlaGlyAspTrpLysThrThrPheThrValAlaGlnAsnGluArgPhe 276

QY 816 GACTTGGTGTTAAACAGAGAATGGAAAGTGTGACCTCACGTTT 860

Db 277 AspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe 291

RESULT 10

JC5248

aryl sulfotransferase (EC 2.8.2.1) HAST2 - human

N;Alternate names: aryl sulfotransferase 1; phenol sulfotransferase; sulfokina

A;Title: Genomic organization and DNA sequences of two human phenol sulfotrans

C;Date: 13-Mar-1997 #sequence revision 25-Apr-1997 #text_change 09-Jul-2004

C;Accession: JC5248; JN0714; JN0479; G01421

R;Dooley, T.P.; Huang, Z.

Biochem. Biophys. Res. Commun. 228, 134-140, 1996

A;Cross-references: UNIPROT:P50225; UNIPARC:UPI000000122E; GB:U71086; NID:g157

R;Zhu, X.; Veronese, M.E.; Bernard, C.C.A.; Sansom, L.N.; McManus, M.E.

A;Reference number: JC5248; MUID:97069665; PMID:8912648

A;Accession: JC5248

A;Status: nucleic acid sequence not shown

A;Molecule type: mRNA

A;Residues: 1-295 <DOO>

Qy	498	GAATTCGCGGAGTTTATGAATGATAAGCTGGCTACGGCTCCTCGTTTGAGCACGTG	555
Db	157	SerPheLeuGluLysPheMetValGlyGluValSerTyrGlySerTrpTyrGlnHisVal	176
Qy	558	CAGGAGTTTCGGGAGCACCCGATGGACTCGAACCTGCTTTTCTCAAAGTATGAAGACATG	617
Db	177	GlnGluTrpGluLeuSerArgThrHisProValLeuTyrLeuPheTyrGluAspMet	196
Qy	618	CATCGGACCTGGTGACATGCTGGACGACGCTGCCAGATCTCTGGGGGTCTCTGTGTAC	677
Db	197	LysGluAsnProLysArgGluIleGlnLysIleLeuGluPheValGlyHisSerLeuPro	216
Qy	678	AAGSCCCAGCTGGGAAGCCCTGACGGAGCAGTCG-----CACCAGCTGGTGGAGCAG	728
Db	217	GluGluThrValAspPheMetValGlnHisThrSerPheLysGluMetLysLysAsnPro	236
Qy	729	TGCTGCAACGCTGAGGCCCTCGCCC-----GTG	755
Db	237	MetThrAsnTyrThrThrValProGlnGluPheMetAspHisSerIleSerProPheMet	256
Qy	756	GCCTGGGGAAGAGTTGGCTGTGGAAGACATCTTCACCGTCTCCATGAATGAGAAAGTTT	815
Db	257	ArgLysGlyMetAlaGlyAspTrpLysThrThrPheThrValAlaGlnAsnGluArgPhe	276
Qy	816	GACTTTGGTGATAAACAGAAGATGGAAAGTGTGACCTTCACGTTT	860
Db	277	AspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe	291

RESULT 11
 JC2523
 aryl sulfotransferase (EC 2.8.2.1) brain isoform - human
 N:Alternate names: phenol sulfotransferase
 C:Species: Homo sapiens (man)
 C;Date: 19-May-1995 #sequence_revision 14-Jul-1995 #text_change 09-Jul-2004

R;Hwang, S.R.; Kohn, A.B.; Hook, V.Y.H.
Biochem. Biophys. Res. Commun. 207, 701-707, 1995
A>Title: Molecular cloning of an isoform of phenol sulfotransferase from human
A;Reference number: JC2523; MUID:95169114; PMID:7864863
A;Accession: JC2523
A;Molecule type: mRNA
A;Residues: 1-295 <HWA>
A;Cross-references: UNIPROT:P50225; UNIPARC:UPI000016A1D1; GB:U26309; NID:g847
A;Experimental source: hippocampus
C;Superfamily: alcohol sulfotransferase
C;Keywords: sulfotransferase
F;70/Active site: Cys #status predicted

Alignment Scores:
Pred. No.: 6,06e-26 Length: 295
Score: 449.00 Matches: 93
Percent Similarity: 54.2% Conservative: 96
Best Local Similarity: 33.8% Mismatches: 108
Query Match: 10.2% Indels: 18
DB: 2 Gaps: 3

US-10-768-158-1 (1-2419) x JC2523 (1-295)

QY	90	GCGGTGCGGCTGCCGCCCTTCTGC	CGCGGAAGATGGAGGATGCCAACTTC	CCGGTG	149
			:::	:::	:::
Db	17	GlyValProLeuIleLysItyrPheAla	GluaLaLeuGlyProLeuGlnSerPhe	GlnAla	36
QY	150	CGCCCCAGCGCGTGTGTGATCGTCA	CCTACCTACC	CCCAAGTCGGCAC	CACAGCTTGTCTGCAGGAG
Db	37	ArgProAspAspLeuLeulleSerThr	TrpProLysSerGlyThrTrpVal	SerGln	56
QY	210	GTGGTCTACTTGTTGGTGAGCCAGG	CGGCGCTGACCCCGATGAGATCGG	CTTGATGA	CATCGAC
Db	57	IleLeuAspMetIleTyrlGlnGly	GlyAspLeuGluLysCysHis	ArgAlaPro	IlePhe
QY	270	GAGCAGTCCC	GGTCTGGAGTACCCACAGCCG	-----	-GGCCTGGACATCATC

A: Experimental source: liver
C: Comment: this enzyme mediates the transfer of SO3- from 3'-phosphoadenosine 5'-phosphocysteine
C: Superfamily: alcohol sulfoltransferase
C: Keywords: sulfoltransferase

Alignment Scores:
Pred. No.: 9 41e-26 Length: 299
Score: 446.50 Matches: 95
Percent Similarity: 55.7% Conservative: 52
Best Local Similarity: 36.0% Mismatches: 98
Query Match: 10.1% Indels: 19
DB: 2 Gaps: 3

US-10-768-158-1 (1-2419) x J55884 (1-299)

```
QY 126 GAGGAGATGCCAACTTCCTCGGTGGCCGCCAGCGAGTGTGGATCGTCACCTACCCCAAG 185
DB 29 GluLysIleGluGluPheGlnSerArgProCysAspIleValIleProThrTyrProLys 48
QY 186 TCCGGCACCAGCTTGTGTCAGGAGGTGTCTACTTGTGTGAGCCAGGGCGCTGACCCCGAT 245
DB 49 SerGlyThrThrTrpLeuSerGluIleValAspMetValLeuAsnAspGlyAsnValGlu 68
QY 246 GAGATCGGCTTGATGAACATGACGAGCAGCTCCGGTCTCTGGAGTACCCACAGCG--- 302
DB 69 LysCysLysArgAspValIleThrSerLysValProMetLeuGluGlnAsnValProGly 88
QY 303 -----GGCTGGACATCATCAAGAACTGACCTCCCGCTCATCAAGAGC 350
DB 89 AlaArgSerGlyValGluLeuLeuLysLysThrProSerProArgIleIleLysThr 108
QY 351 CACCTGCCCTACCGCTTCTGCTGCTGACCTCCACATGGAGACTCCAGGTCTACTAT 410
DB 109 HisLeuProIleAspLeuLeuProLysSerPheThrAspAsnLysCysLysMetIleTyr 128
QY 411 ATGGCTCGCAACCCCAAGGATCTGGTGTCTTATTATCATGCTCCACCGCTCTCTGCGG 470
DB 129 LeuAlaArgAsnGlyLysAspValAlaValSerTyrTyrHisPheAspLeuMetAsnVal 148
QY 471 ACCATGAGCTACCGAGCACCTTTCAAGAAATCTGCGGAGGTATTGAATGAATGAGCTG 530
DB 149 IleGlnProLeuProGlyThrTyrGluGluTyrLeuGluLysPheLeuAlaGlyAsnVal 168
QY 531 GGTACGGCTCTGCTTTCAGACAGCTGAGGAGTCTGGAGCACCGCATGCACTCGAAC 590
DB 169 AlaTyrGlySerTrpPheAspHisValLysSerTrpTrpGluLysArgGluGlyHisPro 188
QY 591 GTGCTTTTCTCAAGTATGAACATGTCATCGGACCTGGTGACGATGGTGAGCAGCTG 650
DB 189 IleLeuPheLeuTyrTyrGluAspLeuLysLysAsnProLysLysGluIleLysLysIle 208
QY 651 GCCAGATTCCTGGGGTGTCTGTGACAGGCCAGCTCGAAGCCCTGACGGAGCACTGC 710
DB 209 AlaAsnPheLeuAspLysThrLeuAspGluHisThrLeuGluArgIleValHisThr 228
QY 711 CACCAGCTGGT-----GACCACTGCTGCACGCTGAGGCCCTGCCC----- 752
DB 229 SerPheGluValMetLysAspAsnProLeuValAsnTyrThrHisLeuProThrGluIle 248
QY 753 -----GTGGCGCGGGGAGAGTGTGGGTGTGGGAGGAGCATC 788
DB 249 MetAspHisSerLysSerProPheMetArgLysGlyValValGlyAspTrpLysAsnTyr 268
QY 789 TTCACGCTCTCCATGAATGAGAGTTTGATCTGGTGTATATAACAGAGATGGAAAGTGT 848
DB 269 PheThrMetThrGlnSerGluLysPheAspAlaIleTyrLysLysLysLeuSerGlyThr 288
QY 849 GACCTCAGCTTT 860
DB 289 ThrLeuGluPhe 292
```

RESULT 14
G02924

phenol sulfoltransferase subunit - crab-eating macaque (fragment)
C: Species: Macaca fascicularis (crab-eating macaque)
C: Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 09-Jul-2004
C: Accession: G02924
R: Ogura, K.
submitted to GenBank, May 1996
A: Reference number: H01975
A: Accession: G02924
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: mRNA
A: Residues: 1-295 <OGU>
A: Cross-references: UNIPROT:P52846; UNIPARC:UPI000013622B; GB:D85514; NID:gl339922; PIDN
C: Superfamily: alcohol sulfoltransferase

Alignment Scores:
Pred. No.: 1.02e-25 Length: 295
Score: 446.00 Matches: 93
Percent Similarity: 54.2% Conservative: 56
Best Local Similarity: 33.8% Mismatches: 108
Query Match: 10.1% Indels: 18
DB: 2 Gaps: 3

US-10-768-158-1 (1-2419) x G02924 (1-295)

```
QY 90 GCGTGGCGCTGCGCCCTCTGCGCGGGAAGATGGAGAGATCGCCAACTTCCCGGTG 149
DB 17 GlyValProLeuIleLysTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheArgAla 36
QY 150 CGGCCACGACGCTGTGGATCGTCACCTACCCCAAGTCCGACACGCTTCTGTCAGAG 209
DB 37 ArgProAspAspLeuLeuIleSerThrTyrProLysSerGlyThrTrpValSerGln 56
QY 210 GTGTCTACTTGTGAGCCAGGCGCTGACCCCGATGAGATCGGCTTGATGAACATCGAC 269
DB 57 IleLeuAspMetIleTyrGlnGlyAspLeuGluLysCysArgAlaProIlePhe 76
QY 270 GAGCAGCTCCCGTCTCGAGTGACCCACAGCCG-----GGCTGGACATCATC 317
DB 77 MetArgValProPheLeuGluPheLysValProGlyIleProSerGlyMetGluThrLeu 96
QY 318 AAGAACTGACCTCTCCCGCTCATCAAGAGCCACCTGCGCTACCGCTTTCTGCCCTCT 377
DB 97 LysAspThrProAlaProArgLeuLeuLysThrHisLeuProLeuAlaLeuProGln 116
QY 378 GACCTCCACATGGAGACTCCCAAGTCTATATGCTCGCTCCCAACCCCAAGATCTGTG 437
DB 117 ThrLeuLeuAspGlnLysValLysValValTyrValAlaArgAsnAlaLysAspValAla 136
QY 438 GTGTCTTATTACGTTCCACCGCTCTCTGCGGACCATGAGCTACCGAGGACCTTTCAA 497
DB 137 ValSerTyrTyrHisPheTyrHisMetAlaLysValHisProGluProGlyThrTrpAsp 156
QY 498 GAATTCCTCCGAGGTATTATGAATGATAGCTGGGCTACGGCTCTCTGTTGAGCAGCTG 557
DB 157 SerPheLeuGluLysPheMetAlaGlyGluValSerTyrGlySerTyrTrpTrpGlnHisVal 176
QY 558 CAGGAGTCTTGGAGCACCGCATGAGCTCGAACAGTCTTTTCTCAAGTATGAACACATG 617
DB 177 GlnGluTrpTrpGluLeuSerHisThrHisProValLeuTyrLeuPheTyrGluAspMet 196
QY 618 CATCGGACCTGGTCAGCATGGTGGAGCAGCTGGCGAGATTCCTGGGGGTGCTGTGAC 677
DB 197 LysGluAsnProLysArgGluIleTrpLysIleLeuGluPheValGlyArgSerLeuPro 216
QY 678 AAGGCCAGCTGGAAGCCCTGACGAGGACATCGC-----CACCAGCTGGTGGACCCAG 728
DB 217 GluGluThrValAspLeuMetValGlnHisThrSerPheLysGluMetLysLysAsnPro 236
QY 729 TCGTCGAACGCTGAGGCCCTGCC-----GTG 755
DB 237 MetAlaAsnTyrThrIleProGlnGluLeuMetAspHisSerIleSerProPheMet 256
QY 755 GCGCGGGGAAGATTGGGCTGTGGAAGGACATCTTTCACCGTCTCCATGAATGAGAAGTTT 815
```

Db 257 ArgLysGlyMetThrGlyAspTrpLysThrPheThrValAlaGlnAsnGluHisPhe 276
QY 816 GACTTGGTGTTATAAACAGAGAGTGGAAAGTGTGACCTCACGTTT 860
Db 277 AspValAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe 291
RESULT 15
S29045
estron sulfotransferase (EC 2.8.2.4) - bovine
N;Alternate names: estrogen sulfotransferase
C;Species: Bos primigenius taurus (cattle)
C;Date: 25-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C;Accession: S18222
R;Nash, A.R.; Glenn, W.K.; Moore, S.S.; Kerr, J.; Thompson, A.R.; Thompson, E.O.P.
Aust. J. Biol. Sci. 41, 507-516, 1988
A;Title: Oestrogen sulfotransferase: molecular cloning and sequencing of cDNA for the bovine
A;Reference number: S29045; MUID: 90147277; PMID: 3271383
A;Accession: S29045
A;Molecule type: mRNA
A;Residues: 1-295 <NAS>
A;Cross-references: UNIPROT:P19217; UNIPARC:UPI0000136223; EMBL:M54942; NID:g163462; PID
A;Experimental source: placenta
R;Adams, J.B.
Biochim. Biophys. Acta 1076, 282-288, 1991
A;Title: Enzymic synthesis of steroid sulphates XVII. On the structure of bovine estrogen
A;Reference number: S13486; MUID: 91152101; PMID: 1900200
A;Accession: S18222
A;Molecule type: protein
A;Residues: 146-160; 206-220; 257-259 <ADA>
A;Cross-references: UNIPARC:UPI0000175826; UNIPARC:UPI0000175827; UNIPARC:UPI0000175828
A;Experimental source: placenta
C;Complex: homodimer
C;Superfamily: alcohol sulfotransferase
C;Keywords: homodimer; phosphoprotein; sulfotransferase
Alignment Scores:
Pred. No.: 1,02e-25 Length: 295
Score: 446.00 Matches: 100
Percent Similarity: 53.8% Conservative: 57
Best Local Similarity: 34.2% Mismatches: 111
Query Match: 10.1% Indels: 24
DB: 2 Gaps: 5
US-10-768-158-1 (1-2419) x S29045 (1-295)
QY 48 AGCACCCCGGGGAGTTCGAGACAAGTACTTCGAGTTCATCGGCTGGCGTCCGCC 107
Db 3 SerSerLysProSerPheSerAspTyrPheGlyLysLeuGlyGlyle-----Pro 19
QY 108 TTCGTCGGCGGGAAGATGAGAGAGATCGCCAAC-----TTCGGGTGCGGCCCAGC 158
Db 20 MetTyrLysLysPheIleGluGlnPheHisAsnValGluGluPheGluAlaArgProAsp 39
QY 159 GAGGTGTGATCTACCTACCCCAAGTCGGCAGCAGTGTCTGTCGAGGAGTGCTTAC 218
Db 40 AspLeuValIleValThrTyrProLysSerGlyThrTyrTrpLeuSerGluIleCys 59
QY 219 TTGCTGAGCCAGGCGCTGACCCCGATGAGATCGGCTTGATGACATCCAGCAGCAGTC 278
Db 60 MetIleTyrAsnAsnGlyAspValGluLysCysLysGluAspValIlePheAsnArgVal 79
QY 279 CCGGTCTCGT-----GAGTACCCACAGCCGGCTGGACATCATCAAGGAACGTG 326
Db 80 ProTyrLeuGluCysSerThrGluHisValMetLysGlyValLysGlnLeuAsnGluMet 99
QY 327 ACCTCTCCCGCTCATCAAGAGCCACCTGCCCTACCGCTTCTGCGCTCTGACCTCCAC 386
Db 100 AlaSerProArgIleValLysSerHisLeuProValLysLeuLeuProValSerPheTrp 119
QY 387 AATGGAGACTCCAAAGTTCATATATATGGTCGCAACCCCAAGGATCTGGTGGTCTTAT 446
Db 120 GluLysAsnCysLysIleIleTyrLeuSerArgAsnAlaLysAspValValSerTyr 139

QY 447 TATCAGTTCCACCGCTCTCTGCGGACCATTAGACTACCGAGGCACCTTTCAAGAAATTCG 506
Db 140 TyrPheLeuIleLeuMetValThrAlaIleProAspProAspSerPheGlnAspPheVal 159
QY 507 CGGAGGTTTATGAATGAAGCTGGCTACGGCTCCTCGTGTGGACACGCTGCAGGAGTTC 566
Db 160 GluLysPheMetCaspGlyGluValProTyrGlySerTrpPheGluHisThrLysSerTrp 179
QY 567 TGGGAGCACCGCATGGACTCGAACGTCGTTTTCTCAAGTATGAAGACATGCATCGGAC 626
Db 180 TrpGluLysSerLysAsnProGlnValLeuPheLeuPheTyrGluAspMetLysGluAsn 199
QY 627 CTGGTGACAGTGTGGAGCAGCTGCCAGATTCCTCGGGGTGCTCTGTGACAAAGCCGAG 686
Db 200 IleArgLysGluValMetLysLeuLeuGluPheLeuGlyArgLysAlaSerAspGluLeu 219
QY 687 CTGGAAGCCCTGACGAGCAGCTGCCAC----- 713
Db 220 ValAspLysIleIleLysHisThrSerPheGlnGluMetLysAsnAsnProSerThrAsn 239
QY 714 -----CAGCTGCTGGACAGCTGCTGCAACGCTGAGGCCCTGCC-----GTGGCGCGGGA 764
Db 240 TyrThrThrLeuProAspGluValMetAsnGlnLysValSerProPheMetArgLysGly 259
QY 765 AGAGTTGGCTGTGAAGACATCTCACCGTCTCCATGAATGAGAAGTTTGTACTTGGTG 824
Db 260 AspValGlyAspTrpLysAsnHisPheThrValAlaLeuAsnGluLysPheAspMethHis 279
QY 825 TATAAACAGAGATGGAAAGTGTGACCTCAGCTTT 860
Db 280 TyrGluGlnGlnMetLysGlySerThrLeuLysPhe 291
RESULT 16
S23399
aryl sulfotransferase (EC 2.8.2.1) - human
C;Species: Homo sapiens (man)
C;Date: 08-May-1995 #sequence_revision 21-Jul-1995 #text_change 18-Jun-1999
C;Accession: S52399
R;Jones, A.L.; Hagen, M.; Coughtrie, M.W.H.; Roberts, R.C.; Glatt, H.R.
submitted to the EMBL Data Library, February 1995
A;Description: Human platelet phenolsulfotransferases: cDNA cloning, stable expression
A;Reference number: S52398
A;Accession: S52399
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-295 <JON>
A;Cross-references: UNIPARC:UPI000016B4D7; EMBL:X84654; NID:g671641; PIDN:CAA59147.1; P1
C;Superfamily: alcohol sulfotransferase
C;Keywords: sulfotransferase
Alignment Scores:
Pred. No.: 1,45e-25 Length: 295
Score: 444.00 Matches: 92
Percent Similarity: 54.2% Conservative: 57
Best Local Similarity: 33.5% Mismatches: 108
Query Match: 10.1% Indels: 18
DB: 2 Gaps: 3
US-10-768-158-1 (1-2419) x S52399 (1-295)
QY 90 GCGTGGCGGCTGCCGCCCTTCTGCGCGGGAAGATGGAGAGATCGCCAACCTCCCGGTG 149
Db 17 GlyValProLeuIleLysTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla 36
QY 150 CGGCCCGGAGCTGTGGATCGTACCTACCCCAAGTCGGCAGCCAGCTTGTGTGAGAG 209
Db 37 ArgProAspAspLeuLeuIleSerThrTyrProLysSerGlyThrTrpValSerGln 56
QY 210 GTGGTCTACTTGGTGGCCGCGCTGACCCCGATGAGATCGGCTTGATGAACATCGAC 269
Db 57 IleLeuAspMetIleTyrGlnGlyGlyAspLeuGluLysCysHisArgAlaProIlePhe 76

```
QY 270 GAGCAGTCCCGTCTCGAGTACCCACAGCGG-----GGCTGGACATCATC 317
Db 270 GAGCAGTCCCGTCTCGAGTACCCACAGCGG-----GGCTGGACATCATC 317
QY 77 MetArgValProPheLeuGluPheValProGlyIleProSerGlyMetGluThrLeu 96
Db 77 MetArgValProPheLeuGluPheValProGlyIleProSerGlyMetGluThrLeu 96
QY 318 AAGGAACGACCTCTCCCGCTCATCAAGAGCCACCTCCGCTTCTGCGCTCT 377
Db 318 AAGGAACGACCTCTCCCGCTCATCAAGAGCCACCTCCGCTTCTGCGCTCT 377
QY 97 LysAspThrProAlaProArgLeuLeuLysThrHisLeuProLeuAlaLeuProGln 116
Db 97 LysAspThrProAlaProArgLeuLeuLysThrHisLeuProLeuAlaLeuProGln 116
QY 378 GACCTCCCAATGGAGACTCCCAAGGTCTATATGGTTCGCAACCCCAAGGATCGTG 437
Db 378 GACCTCCCAATGGAGACTCCCAAGGTCTATATGGTTCGCAACCCCAAGGATCGTG 437
QY 117 ThrLeuLeuAspGlnLysValLysValValTyrValAlaArgAsnAlaLysValAla 136
Db 117 ThrLeuLeuAspGlnLysValLysValValTyrValAlaArgAsnAlaLysValAla 136
QY 438 GTGCTTATATCAGTTCCACCGCTCTCTGCGGACCATGAGTACCGAGGACCTTCAA 497
Db 438 GTGCTTATATCAGTTCCACCGCTCTCTGCGGACCATGAGTACCGAGGACCTTCAA 497
QY 137 ValSerTyrTyrHisPheTyrHisMetThrLysValHisProGluProGlyThrTrpAsp 156
Db 137 ValSerTyrTyrHisPheTyrHisMetThrLysValHisProGluProGlyThrTrpAsp 156
QY 498 GAATTCCTGCGGAGGTTTATGATAGCTGGGTACGGCTCTGTTGTTGAGCAGCTG 557
Db 498 GAATTCCTGCGGAGGTTTATGATAGCTGGGTACGGCTCTGTTGTTGAGCAGCTG 557
QY 157 SerPheLeuGluLysPheMetValGlyLysValSerTyrGlySerTyrGlnHisVal 176
Db 157 SerPheLeuGluLysPheMetValGlyLysValSerTyrGlySerTyrGlnHisVal 176
QY 558 CAGGAGTCTCGGAGCACCAGCATCGACTCGAACGTCGTTTTCTCAAGTATGAAGACATG 617
Db 558 CAGGAGTCTCGGAGCACCAGCATCGACTCGAACGTCGTTTTCTCAAGTATGAAGACATG 617
QY 177 GlnGluTyrTrpGlyLeuSerArgThrHisProValLeuTyrLeuPheTyrGluAspMet 196
Db 177 GlnGluTyrTrpGlyLeuSerArgThrHisProValLeuTyrLeuPheTyrGluAspMet 196
QY 618 CATCGGACCTGGTACGATGGTGGAGCAGCTGGCCAGATTCTCGGGGTGTCCTGTGAC 677
Db 618 CATCGGACCTGGTACGATGGTGGAGCAGCTGGCCAGATTCTCGGGGTGTCCTGTGAC 677
QY 197 LysGluAsnProLysArgGluLeuGlnLysIleLeuGluPheValGlyHisSerLeuPro 216
Db 197 LysGluAsnProLysArgGluLeuGlnLysIleLeuGluPheValGlyHisSerLeuPro 216
QY 678 AAGGCCACCTGGAGCCCTGACGGAGCAGCTGC-----CACCAGCTGGTGACACAG 728
Db 678 AAGGCCACCTGGAGCCCTGACGGAGCAGCTGC-----CACCAGCTGGTGACACAG 728
QY 217 GluGluThrValAspPheMetValGlnHisThrSerPheLysGluMetLysAsnPro 236
Db 217 GluGluThrValAspPheMetValGlnHisThrSerPheLysGluMetLysAsnPro 236
QY 729 TGCTGCAACGCTGAGCGCTCGGCC-----GTG 755
Db 729 TGCTGCAACGCTGAGCGCTCGGCC-----GTG 755
QY 237 MetThrAsnTyrThrThrValProGlnGluPheMetAspHisSerIleSerProPheMet 256
Db 237 MetThrAsnTyrThrThrValProGlnGluPheMetAspHisSerIleSerProPheMet 256
QY 756 GCGCGGGGAGAGTTGGGTGGAGGACATCTTCCCGTCTCCATGATGAGAAGTTT 815
Db 756 GCGCGGGGAGAGTTGGGTGGAGGACATCTTCCCGTCTCCATGATGAGAAGTTT 815
QY 257 ArgLysGlyMetAlaGlyAspTrpLysThrPheThrPheValAlaGlnAsnGluArgPhe 276
Db 257 ArgLysGlyMetAlaGlyAspTrpLysThrPheThrPheValAlaGlnAsnGluArgPhe 276
QY 816 GACTTGGTCTATAACAGAGATGGGAAGTGTGACCTCACGTTT 860
Db 816 GACTTGGTCTATAACAGAGATGGGAAGTGTGACCTCACGTTT 860
QY 277 AspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe 291
Db 277 AspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe 291

RESULT 17
JCS585
thyroid hormone sulfotransferase (EC 2.8.2.-) B2 - human
C;Species: Homo sapiens (man)
C;Date: 11-Mar-1998 #sequence_revision 11-Mar-1998 #text_change 09-Jul-2004
C;Accession: JCS585
R;Fujita, K.; Nagata, K.; Ozawa, S.; Sasano, H.; Yamazoe, Y.
J. Biochem. 122, 1052-1061, 1997
A;Title: Molecular cloning and characterization of rat ST1B1 and human ST1B2 cDNAs, encoding
A;Reference number: JCS584; MUID:98104061; PMID:9443824
A;Accession: JCS585
A;Molecule type: mRNA
A;Residues: 1-296 <FUJ>
A;Cross-references: UNIPROT:O43704; UNIPARC:UPI0000073458; DDBJ:D89479; NID:g2826145; PIR:
A;Experimental source: liver
C;Comment: This enzyme mediates the transfer of fSOS- from 3'-phosphoadenosine 5'-phospho
C;Superfamily: alcohol sulfotransferase
C;Keywords: sulfotransferase

Alignment Scores:
Prod. No.: 2,68e-25 Length: 296
Score: 440.50 Matches: 94
Percent Similarity: 53.4% Conservative: 54
Best Local Similarity: 33.9% Mismatches: 110
Query Match: 10.0% Indels: 19
DB: 2 Gaps: 3

US-10-768-158-1 (1-2419) x JCS585 (1-296)
```

```
QY 87 CATGGCTGGCGTCCCGCCCTTCTGCGCGGGAAGATGGAGGAGATCGCAACTTCCG 146
Db 87 CATGGCTGGCGTCCCGCCCTTCTGCGCGGGAAGATGGAGGAGATCGCAACTTCCG 146
QY 16 HisGlyTyrProMetThrCysAlaPheAlaSerAsnTrpGluLysIleGluGlnPheHis 35
Db 16 HisGlyTyrProMetThrCysAlaPheAlaSerAsnTrpGluLysIleGluGlnPheHis 35
QY 147 GTGGCGGCCACGCGTGTGGATCGTACCTTACCCCAAGTCCGGCACCAGCTTGTGCGAG 206
Db 147 GTGGCGGCCACGCGTGTGGATCGTACCTTACCCCAAGTCCGGCACCAGCTTGTGCGAG 206
QY 36 SerArgProAspAspIleValIleAlaThrTyrProLysSerGlyThrThrTrpValSer 55
Db 36 SerArgProAspAspIleValIleAlaThrTyrProLysSerGlyThrThrTrpValSer 55
QY 207 GAGTGTGCTACTTGGTGAGCCAGGGCGCTGACCCCGATGAGATCGGCTTGAACATC 266
Db 207 GAGTGTGCTACTTGGTGAGCCAGGGCGCTGACCCCGATGAGATCGGCTTGAACATC 266
QY 56 GluIleIleAspMetIleLeuAsnAspGlyIleGluLysCysLysArgGlyPheIle 75
Db 56 GluIleIleAspMetIleLeuAsnAspGlyIleGluLysCysLysArgGlyPheIle 75
QY 267 GACGAGCAGCTCCCGGCTCTGAGTACCCACAGCGGCGCTG-----GAC 311
Db 267 GACGAGCAGCTCCCGGCTCTGAGTACCCACAGCGGCGCTG-----GAC 311
QY 76 ThrGluLysValProMetLeuGluMetThrLeuProGlyLeuArgThrSerGlyIleGlu 95
Db 76 ThrGluLysValProMetLeuGluMetThrLeuProGlyLeuArgThrSerGlyIleGlu 95
QY 312 ATCATCAAGAACTGACCTCTCCCGCTCATCAAGAGCACCCTGCGCTACCGCTTTCTG 371
Db 312 ATCATCAAGAACTGACCTCTCCCGCTCATCAAGAGCACCCTGCGCTACCGCTTTCTG 371
QY 96 GlnLeuGluLysAsnProSerProArgIleValLysThrHisLeuProThrAspLeuLeu 115
Db 96 GlnLeuGluLysAsnProSerProArgIleValLysThrHisLeuProThrAspLeuLeu 115
QY 372 CCCTCTGACCTCCCAATGGAGACTCCCAAGGTCTATATGGTTCGCGCAACCCCAAGAT 431
Db 372 CCCTCTGACCTCCCAATGGAGACTCCCAAGGTCTATATGGTTCGCGCAACCCCAAGAT 431
QY 116 ProLysSerPheTrpGluAsnAsnCysLysMetIleTyrLeuAlaArgAsnAlaLysAsp 135
Db 116 ProLysSerPheTrpGluAsnAsnCysLysMetIleTyrLeuAlaArgAsnAlaLysAsp 135
QY 432 CTGGTGGTGTCTTATTATCAGTTCCACGCTCTCTGCGGACCATGAGCTACCGAGGCACC 491
Db 432 CTGGTGGTGTCTTATTATCAGTTCCACGCTCTCTGCGGACCATGAGCTACCGAGGCACC 491
QY 136 ValSerValSerTyrTyrHisPheAspLeuMetAsnLeuGlnProPheProGlyThr 155
Db 136 ValSerValSerTyrTyrHisPheAspLeuMetAsnLeuGlnProPheProGlyThr 155
QY 492 TTTCAGAAATCTTGGCGGAGGTTTATGAATGATGAAGCTGGGCTACGGCTCTCTGTTGAG 551
Db 492 TTTCAGAAATCTTGGCGGAGGTTTATGAATGATGAAGCTGGGCTACGGCTCTCTGTTGAG 551
QY 156 TrpGluGluTyrLeuGluLysPheLeuThrGlyLysValAlaTyrGlySerTrpPheThr 175
Db 156 TrpGluGluTyrLeuGluLysPheLeuThrGlyLysValAlaTyrGlySerTrpPheThr 175
QY 552 CACGTGACGAGAGTTCTGGAGCACCAGCATGGTGGAGCAGCTGGCAGATTCCTGGGGTGTC 611
Db 552 CACGTGACGAGAGTTCTGGAGCACCAGCATGGTGGAGCAGCTGGCAGATTCCTGGGGTGTC 611
QY 176 HisValLysAsnTrpTrpLysLysGluGlyHisProIleLeuPheLeuTyrTrpGlu 195
Db 176 HisValLysAsnTrpTrpLysLysGluGlyHisProIleLeuPheLeuTyrTrpGlu 195
QY 612 GACATGCATCGGACCTGGTACGATGGTGGAGCAGCTGGCAGATTCCTGGGGTGTC 671
Db 612 GACATGCATCGGACCTGGTACGATGGTGGAGCAGCTGGCAGATTCCTGGGGTGTC 671
QY 196 AspMetLysGluAsnProLysGluGluLysIleArgPheLeuGluLysAsn 215
Db 196 AspMetLysGluAsnProLysGluGluLysIleArgPheLeuGluLysAsn 215
QY 672 TGTGACAGGCCAGCTGGAAGCCCTGACGGAGCAGCTGCCACACAGCTGGT----- 722
Db 672 TGTGACAGGCCAGCTGGAAGCCCTGACGGAGCAGCTGCCACACAGCTGGT----- 722
QY 216 LeuAsnAspGluIleLeuAspArgIleIleHisHisThrSerPheGluValMetLysAsp 235
Db 216 LeuAsnAspGluIleLeuAspArgIleIleHisHisThrSerPheGluValMetLysAsp 235
QY 723 GACCAGTGTGCAACGCTGAGCGCTCGCC----- 752
Db 723 GACCAGTGTGCAACGCTGAGCGCTCGCC----- 752
QY 236 AsnProLeuValAsnTyrThrHisLeuProThrThrValMetAspHisSerLysSerPro 255
Db 236 AsnProLeuValAsnTyrThrHisLeuProThrThrValMetAspHisSerLysSerPro 255
QY 753 ---GTGGCGCGGGAAGAGTTGGGCTGTGGAGGACATCTTCCCGTCTCCATGATGAG 809
Db 753 ---GTGGCGCGGGAAGAGTTGGGCTGTGGAGGACATCTTCCCGTCTCCATGATGAG 809
QY 256 PheMetArgLysGlyThrAlaGlyAspTrpLysAsnTyrPheThrValAlaGlnAsnGlu 275
Db 256 PheMetArgLysGlyThrAlaGlyAspTrpLysAsnTyrPheThrValAlaGlnAsnGlu 275
QY 810 AAGTTTGAAGTCTGTATATAACAGAGATGGGAAAGTGTGACCTCACGTTT 860
Db 810 AAGTTTGAAGTCTGTATATAACAGAGATGGGAAAGTGTGACCTCACGTTT 860
QY 276 LysPheAspAlaIleTyrGluThrGluMetSerLysThrAlaLeuGlnPhe 292
Db 276 LysPheAspAlaIleTyrGluThrGluMetSerLysThrAlaLeuGlnPhe 292

RESULT 18
I73679
estrogen sulfotransferase isoform 6 - rat
C;Species: Rattus sp. (rat)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 18-Jun-1999
C;Accession: I73679
R;Falany, J.L.; Krasnykh, V.; Mikheeva, G.; Falany, C.N.
J. Steroid Biochem. Mol. Biol. 52, 35-44, 1995
A;Title: Isolation and expression of an isoform of rat estrogen sulfotransferase.
A;Reference number: I56606; MUID:95161323; PMID:7857871
A;Accession: I73679
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-295 <RES>
A;Cross-references: UNIPARC:UPI0000136222; GB:S76490; NID:g913356; PIDN:AAB33442.1; PID
C;Superfamily: alcohol sulfotransferase
```

Alignment Scores:				
Pred. No.:	7,01e-25	Length:	295	
Score:	435.00	Matches:	93	
Percent Similarity:	53.1%	Conservative:	62	
Best Local Similarity:	31.8%	Mismatches:	105	
Query Match:	9.9%	Indels:	32	
DB:	2	Gaps:	4	
US-10-768-158-1 (1-2419) x I73679 (1-295)				
QY	60	GAGTTCGACGACGAGTACTTCGAGTTCATGGCGTGGCGTGGCGGCTTCGCGCGGG	119	
DB	7	GlutyrTyrGluValPheGlyAspPheHisGlyValLeuValAspLysLeuPheThrLys	26	
QY	120	AAGATGGAGGAGATCCCAACTTCGCGTGGCGGCCGAGCGAGTGTGGATCGTCACTAC	179	
DB	27	TyrTrpGluAspValGluThrPheSerAlaAsgProAspAspLeuValValThrTyr	46	
QY	180	CCCAAGTCCGGACCACTTGTCGACGAGGTGGTCTACTTGGTAGCCAGGCGGTGAC	239	
DB	47	ProLysSerGlySerThrTrpIleGlyGluIleValAspMetIleTyrLysGluGlyAsp	66	
QY	240	CCCGATGAGATCGGCTTGATGAACATCGACGAGCAGCTCCCGTCTCGAGTACCCACAG	299	
DB	67	ValGluLysCysLysGluAspAlaIlePheAsnArgIleProTyrLeuGluCysArgAsn	86	
QY	300	CCG-----GGCTGGACATCATCAAGGAACCTGACCTCTCCCGCTCATCAAG	347	
DB	87	GluAspLeuIleAsnGlyIleLysGlnLeuLysGluLysGluSerProArgIleValLys	106	
QY	348	AGCCACCTGCGCTACCGCTTCTGCGCTCTGACCTCCCAATGGAGACTCCAAGGTGATC	407	
DB	107	ThrHisLeuProAlaLysLeuLeuProAlaSerPheTrpGluLysAsnCysLysIleIle	126	
QY	408	TATATGGCTCGCAACCCCAAGGATCTGGTGGTCTTATTATCAGTTCCACCGCTCTCG	467	
DB	127	TyrLeuCysArgAsnAlaLysAspValValSerTyrTyrPhePheLeuIleIle	146	
QY	468	CGACCATGAGCTACCGAGGCACCTTCAAGAATCTCGCGGAGGTTTATGAATGAAG	527	
DB	147	LysSerTyrProAsnProLysSerPheSerGluPheValGluLysPheMetGluGln	166	
QY	528	CTGGCTACGGCTCTCTGTTTTCAGCAGCTGAGGAGTCTGGAGCAGCCGATGAGCTCG	587	
DB	167	ValProTyrGlySerTrpTyrAspHisValLysSerTrpTrpGluLysSerLysAsnSer	186	
QY	588	AACGTGCTTTTCTCAAGTATCAAGACATGCATCGGCACCTCGTGACGATGTGGAGCAG	647	
DB	187	ArgValLeuPheMetPheTyrGluAspMetLysGluAspIleArgGluValValLys	206	
QY	648	CTGGCCAGATTCCTGGGGGTCTCTGTGACAAAGGCCAGCTCGAAGCCCTGACGGAGCAC	707	
DB	207	LeuIleGluPheLeu-----GluArgAspProLeuAlaGluLeu	219	
QY	708	TCCACACAGCTGGTG-----GACCAGTGTGCAAC	737	
DB	220	ValAspLysIleIleGlnHisThrPheGlnGluMetLysAsnAsnProCysThrAsn	239	
QY	738	GCTGAGGCCCTCCCT-----GTGGCCGGGA	764	
DB	240	TyrSerMetLeuProGluThrMetIleAspLeuLysValSerProPheMetArgLysGly	259	
QY	765	AGATGTGGCTGTGAAGGACATCTTACCGTCTCCATGAATGAGAAGTTTACACTGGTG	824	
DB	260	IleValGlyAspTrpArgAsnHisPheProGluAlaLeuArgGluArgPheGluHis	279	
QY	825	TATAACAGACAGATGGAAAGTGTGACCTCAGCTTT	860	
DB	280	TyrGlnArgHisMetLysAspCysProValThrPhe	291	
RESULT 19				
JC2229 estrogen sulfotransferase (EC 2.8.2.-) - human				

C:Species: Homo sapiens (man)				
C>Date:	28-Aug-1985	#sequence	revision	07-Oct-1994 #text_change
C:Accession:	JC2229; I38249; I56607			
R:Aksoy,	I.A.; Wood,	T.C.; Weinshilboum,	R.	
Biochem.	Biophys. Res. Commun.	200	1621-1629,	1994
A:Title:	Human liver estrogen sulfotransferase: identification by cDNA cloning and expres			
A:Reference number:	JC2229; MUID:94242031; PMID:8185618			
A:Accession:	JC2229			
A:Molecule type:	mRNA			
A:Residues:	1-294 <AKS>			
A:Cross-references:	UNIPROT:P49888; UNIPARC:UPI00000623C8; GB:U08098; NID:9488282; PIDN:			
A:Experimental source:	liver			
R:Her,	C.; Aksoy,	I.A.; Kimura,	S.; Brandriff,	B.F.; Wasmuth,
Genomics	29, 16-23,	1995		
A:Title:	Human estrogen sulfotransferase gene (STE): Cloning, structure and chromosomal			
A:Reference number:	I38249; MUID:96079087; PMID:8530066			
A:Accession:	I38249			
A>Status:	preliminary			
A:Molecule type:	DNA			
A:Residues:	1-294 <RES>			
A:Cross-references:	UNIPARC:UPI00000623C8; EMBL:U20521; NID:g1079511; PIDN:AAC50286.1; PID:			
R:Palany,	C.N.; Krasnykh,	V.; Falany,	J.L.	
J. Steroid	Biochem. Mol. Biol.	52, 529-539,	1995	
A:Title:	Bacterial expression and characterization of a cDNA for human liver estrogen s			
A:Reference number:	I56607; MUID:95298714; PMID:7779757			
A:Accession:	I56607			
A>Status:	preliminary; translated from GB/EMBL/DDBJ			
A:Molecule type:	mRNA			
A:Residues:	1-294 <RES>			
A:Cross-references:	UNIPARC:UPI00000623C8; GB:S77383; NID:g998887; PIDN:AAB34601.1; PID:			
C:Comment:	This enzyme catalyzes the sulfate conjugation of estrogen.			
C:Genetics:				
A:Gene:	GDB:STE			
A:Cross-references:	GDB:342104; OMIM:600043			
A:Map position:	16pter-16qter			
A:Introns:	49/1, 91/1, 123/3, 166/1, 197/3, 258/1			
C:Superfamily:	alcohol sulfotransferase			
C:Keywords:	sulfotransferase			
Alignment Scores:				
Pred. No.:	8,35e-25	Length:	294	
Score:	434.00	Matches:	95	
Percent Similarity:	53.0%	Conservative:	56	
Best Local Similarity:	33.3%	Mismatches:	116	
Query Match:	9.9%	Indels:	18	
DB:	2	Gaps:	3	
US-10-768-158-1 (1-2419) x JC2229 (1-294)				
QY	60	GAGTTCGACGACGAGTACTTCGAGTTCATGGCGTGGCGTGGCGGCTTCGCGCGGG	119	
DB	6	AspTyrTyrGluLysPheGluGluValHisGlyIleLeuMetTyrLysAspPheValLys	25	
QY	120	AAGATGGAGGAGATCGCCAACTTCCCGTGGCGGCCGAGCGTGTGGATCGTCACTAC	179	
DB	26	TyrTrpAspAsnValGluAlaPheGlnAlaArgProAspAspLeuValIleAlaThrTyr	45	
QY	180	CCCAAGTCCGACGAGCTTGTCTGAGGAGTGTGTCTACTTGGTGGAGCCAGGCGGCTGAC	239	
DB	46	ProLysSerGlyThrTrpValSerGluIleValIleMetIleTyrLysGluGlyAsp	65	
QY	240	CCCGATGAGATCGGCTTGATGAACATCGACGAGCAGTCCCGGTCTCGAGTACCCACAG	299	
DB	66	ValGluLysCysLysGluAspValIlePheAsnArgIleProPheLeuGluCysArgLys	85	
QY	300	CCG-----GGCTGGACATCATCAAGGAACCTGACCTCCCGGTCTCGAGTACCCACAG	347	
DB	86	GluAsnLeuMetAsnGlyValLysGlnLeuAspGluMetAsnSerProArgIleValLys	105	
QY	348	AGCCACCTGCGCTACCGCTTCTGCGCTCTGACCTCCCAATGGAGACTCCAAGGTGATC	407	
DB	106	ThrHisLeuProGluLeuLeuProAlaSerPheTrpGluLysAspCysLysIleIle	125	

```
QY 408 TATATGGCTCGCAACCCCAAGGACTGTGGTGTCTTATTATCAGTTCACCGCTCTCTG 467
    |||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 126 TyrLeuCyAsrGAsnAlaYsAspValAlaValSerPheTyrTyrPhePheLeuMetVal 145
    |||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 468 CGGACCATCAGCTACCGAGGACCTTTCAAGAAATTCGCGGAGGTTCATGAATGAATAG 527
    |||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 146 AlaGlyHisProAsnProGlySerPheProGluPheValGluLysPheMetGlnGlyGln 165
    |||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 528 CTGGGCTACGGCTCTCTGTTTACGACGTGCAGGAGTTCCTGGGAGCAGCGCATGACTCG 587
    |||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 166 ValProTyrGlySerTrpTyrIleHisValLysSerTrpTrpGluLysGlySerPro 185
    |||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 588 AACGTCCTTTTCTCAAGTATCAAGACATCATCGGACCTCGTGCAGCATGTGCGAGCAG 647
    |||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 186 ArgValLeuPheLeuPheTyrGluAspLeuLysGluAspIleArgLysGluValIleLys 205
    |||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 648 CTGGCCAGATCTCTGGGGTGTCTCTGACAAAGCCGACCTGGAAGCCCTGACGGAGCAC 707
    |||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 206 LeuIleHisPheLeuGluArgLysProSerGluGluLeuValAspArgIleIleHis 225
    |||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 708 TGCCAC-----CAGCTGTGGACGAC 728
    |||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 226 ThrSerPheGlnGluMetLysAsnProSerThrAsnTyrThrLeuProAspGlu 245
    |||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 729 TCGTCAACGCTGAGGCCCTGCC---GTGGCCGGGAGAGTTGGGCTGTGGAAGGAC 785
    |||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 246 IleMetAsnGlnLysLeuSerProPheMetArgLysGlyLeuThrGlyAspTrpLysAsn 265
    |||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 786 ATCTTCACCGCTCCATGAATGAGAAGTTTGATTTGCTGTATAAACAAGATGGGAAG 845
    |||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 266 HisPheThrValAlaLeuAsnGluLysPheAspLysHisTyrGluGlnMetLysGlu 285
    |||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 846 TGTGACCTCAGCTTT 860
    |||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 286 SerThrLeuLysPhe 290
    |||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 20
JE0186
sulfoftransferase (EC 2.8.-.-) SULT1B1 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 21-Aug-1998 #sequence_revision 21-Aug-1998 #text_change 31-Dec-2004
C:Accession: JE0186
R:Saeki, Y.; Sakakibara, Y.; Araki, Y.; Yanagisawa, K.; Suiko, M.; Nakajima, H.; Liu, M.
J. Biochem. 124, 55-64, 1998
A:Title: Molecular cloning, expression, and characterization of a novel mouse liver SULT
A:Reference number: JE0186; MUID:98309850; PMID:9644246
A:Accession: JE0186
A:Molecule type: mRNA
A:Residues: 1-299 <SAE>
A:Cross-references: UNIPROT:Q9QW67; UNIPROT:Q9Z2T0; UNIPARC:UPI000017582F; GB:AF022894
A:Experimental source: liver
C:Comment: This enzyme plays a role in the homeostatic regulation of 3,4-dihydroxyphenyl
C:Keywords: transferase

Alignment Scores:
Pred. No.: 1,09e-24 Length: 299
Score: 432.50 Matches: 94
Percent Similarity: 53.7% Conservative: 50
Best Local Similarity: 35.1% Mismatches: 97
Query Match: 9.8% Indels: 27
DB: 2 Gaps: 4

US-10-768-158-1 (1-2419) x JE0186 (1-299)
```

```
Db 69 LysCysLysArgAspValIleThrSerLysValProMetLeuGluLeuSerValProGly 88
    |||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 306 CTG-----GACATCATCAAGAACTGACCTCTCCCGCGCTCATCAAGAGC 350
    |||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 89 IleArgIleSerGlyValGluLeuLysLysThrProSerProArgIleIleLysThr 108
    |||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 351 CACCTGCCCTACCGCTTCTGCCCTCTGACCTCCACAATGGAGACTCCCAAGTCTATCTAT 410
    |||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 109 HisLeuProIleAspLeuLeuProLysSerPheTrpGluAsnLysCysLeuMetIleLys 128
    |||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 411 ATGGCTCCCAACCCCAAGGATCTGTGTCTTATTATCAGTTCCTCCACCGCTCTCTGCGG 470
    |||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 129 LeuAlaArgAsnGlyLysAspValAlaValSerTyrTyrHisPheAspLeuMetAsnSer 148
    |||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 471 ACCATGAGCTACCGAGGACCTTTCAAGAAATTCGCGGAGGTTCATGAATGAATAGCTG 530
    |||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 149 IleAsnProLeuProGlyThrTrpGluGluTyrLeuGluLysPheLeuAlaGlyAsnVal 168
    |||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 531 GGCTACGGCTCTGGTTTGACAGCTGCGAGTCTCTGGGAGCACCGCATGGACTCGAAC 590
    |||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 169 AlaTyrGlySerTrpPheGluHisValLysSerTrpTrpGluLysArgGluHisPro 188
    |||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 591 GTGCTTTTCTCAAGTATGAAGACATCATCGGACCTGTGTGTCAGCATGTGTGGAGCAGCTG 650
    |||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 189 LeuLeuTyrLeuTyrTyrGluLeuLysGlnAsnProLysLysGluIleLysLysIle 208
    |||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 651 GCCAGATTCTCGGGGTGTCTGTGACAAGCCCGAGCTGGAAGCCCTCAGCGAGCAGCTGC 710
    |||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 209 AlaSerPheLeuAspLysThrLeuAspGluLeuAlaLeuAspArgIleValHisThr 228
    |||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 711 -----CACCAGCTGTGGACCACTGTGCAACGCTGAGGCCCTG 749
    |||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 229 SerPheGluMetMetLysGluAsnProLeuVal-----AsnTyrThrHisLeu 244
    |||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 750 CCCGTGGGC-----CGGGGAAGAGTTGGGCTG 776
    |||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 245 ProThrAlaMetMetAspHisSerLysSerProPheMetArgLysGlyLeuValGlyAsp 264
    |||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 777 TGAAGAGACATCTTCACCGCTCCATGAATGAGAAGTTTGACTGGTGTATAACAGAGAG 836
    |||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 265 TrpLysAsnTyrPheThrMetThrGlnThrGluGlnPheAspAlaValTyrLysLysLys 284
    |||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 837 ATGGGAAGTGTGACCTCAGCTTT 860
    |||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 285 MetSerGlyThrThrLeuGluPhe 292
    |||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 21
I53296
testis-specific estrogen sulfotransferase - mouse
C:Species: Mus sp. (mouse)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 18-Jun-1999
C:Accession: I53296
R:Song, W.C.; Moore, R.; McLachlan, J.A.; Negishi, M.
Endocrinology 136, 2477-2484, 1995
A:Title: Molecular characterization of a testis-specific estrogen sulfotransferase and a
A:Reference number: I53296; MUID:95269690; PMID:7750469
A:Accession: I53296
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-295 <RES>
A:Cross-references: UNIPARC:UPI000014D493; GB:S78182; NID:G999090; PIDN:AAB34320.1; PID
C:Superfamily: alcohol sulfotransferase

Alignment Scores:
Pred. No.: 1.68e-24 Length: 295
Score: 430.00 Matches: 94
Percent Similarity: 53.0% Conservative: 57
Best Local Similarity: 33.0% Mismatches: 116
Query Match: 9.8% Indels: 18
DB: 2 Gaps: 3

US-10-768-158-1 (1-2419) x I53296 (1-295)
```

```
QY 60 GAGTTCGAGAGCAAGTACTTCGAGTTCCATGGCGTGGCGCTGCCGCTTCTGCCGCGG 119
DB 7 GlutyrTyrGluValPheGlyGluPheArgGlyValLeuMetAspLysArgPheThrLys 26
QY 120 AAGATCGAGGAGATCCCAACTTCCCGGTCCGCGCCAGCGAGCTGTGGATCGTCACTAC 179
DB 27 TyrTrpGluAspValGluMetPheLeuAlaArgProAspAspLeuValIleAlaThrTyr 46
QY 180 CCCAAGTCGCGCACCAAGCTTGTGCGAGGAGTGTCTACTTGTGAGCAGCGCGCTGAC 239
DB 47 ProLysSerGlyThrTrpIleSerGluValValTyrMetIleTyrLysGluGlyAsp 66
QY 240 CCGGATGAGATCGGCTTGATGAACATCGACGACAGCTCCCGTCTCTGGAGTACCCACAG 299
DB 67 ValGluLysCysLysGluAspAlaIlePheAsnArgIleProTyrLeuGluCysArgAsn 86
QY 300 CCG-----GGCCTGGACATCATCAAGAACTGACCTCTCCCGCTCATCAAG 347
DB 87 GluAspLeuIleAsnGlyIleLysGlnLeuLysGluSerProArgIleValLys 106
QY 348 AGCCACCTCGCTACCGCTTCTGCGCTTCTGCGCTTCTGCGCTTCTGCGCTTCTGCGCT 407
DB 167 ValProTyrGlySerTrpTyrAspHisValLysAlaTrpTrpGluLysSerLysAsnSer 186
QY 588 AACGTCCTTTTCTCAAGTATGAAGACATCGCATCGGACCTCGTGCAGTGTGGAGCAG 647
DB 187 ArgValLeuPheMetPheTyrGluAspMetLysGluAspIleArgArgGluValValLys 206
QY 648 CTGGCCAGATTCCTGGGGGTCTGCTGTGACAGGCCAGCTCGAAGCCCTGAGGAGCAG 707
DB 207 LeuIleGluPheLeuGluArgLysProSerAlaGluLeuValAspArgIleIleGlnHis 226
QY 846 TGTGACCTCAGCTTT 860
DB 287 CysThrValLysPhe 291
RESULT 22
156606
estrogen sulfotransferase isoform 3 - rat
C:Species: Rattus sp. (rat)
C:Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 18-Jun-1999
C:Accession: 156606
R:Falany, J.L.; Krasnykh, V.; Mikheeva, G.; Falany, C.N.
J. Steroid Biochem. Mol. Biol. 52, 35-44, 1995
A>Title: Isolation and expression of an isoform of rat estrogen sulfotransferase.
A:Reference number: 156606; M01D:195161323; PMID:7857871
A:Accession: 156606
A>Status: preliminary; translated from GB/EMBL/DBJ
```

A:Molecule type: mRNA
A:Residues: 1-295 <RES>
A:Cross-references: UNIPARC:UPI0000170CB4; GB:S76489; NID:g913354; PIDN:AAB33441.1; PID
C:Superfamily: alcohol sulfotransferase

Alignment Scores:
Pred. No.: 1.68e-24 Length: 295
Score: 430.00 Matches: 92
Percent Similarity: 53.3% Conservative: 60
Best Local Similarity: 32.3% Mismatches: 115
Query Match: 9.8% Indels: 18
DB: 2 Gaps: 3

US-10-768-158-1 (1-2419) x 156606 (1-295)

```
QY 60 GAGTTCGAGAGCAAGTACTTCGAGTTCCATGGCGTGGCGCTGCCGCTTCTGCCGCGG 119
DB 7 GlutyrTyrAspValPheGlyAspPheHisGlyPheLeuMetAspLysArgPheThrLys 26
QY 120 AAGATCGAGGAGATCGCCAACTTCCCGGTCCGCGCCAGCGAGCTGTGGATCGTCACTAC 179
DB 27 TyrTrpGluAspValGluThrPheLeuAlaArgProAspAspLeuLeuIleValThrTyr 46
QY 180 CCCAAGTCGCGCACCAAGCTTGTGCGAGGAGTGTCTACTTGTGAGCAGCGCGCTGAC 239
DB 47 ProLysSerGlySerThrTrpIleSerGluIleValAspMetIleTyrLysGluGlyAsp 66
QY 240 CCGGATGAGATCGGCTTGATGAACATCGACGAGAGCTCCCGTCTCTGGAGTACCCACAG 299
DB 67 ValGluLysCysLysGluAspAlaLeuPheAsnArgIleProAspLeuGluCysArgAsn 86
QY 300 CCG-----GGCCTGGACATCATCAAGAACTGACCTCTCCCGCTCATCAAG 347
DB 87 GluAspLeuIleAsnGlyIleLysGlnLeuLysGluSerProArgIleValLys 106
QY 348 AGCCACCTCGCTACCGCTTCTGCGCTTCTGCGCTTCTGCGCTTCTGCGCTTCTGCGCT 407
DB 107 ThrHisLeuProAlaLysLeuLeuProAlaSerPheTrpGluLysAsnCysLysIleIle 126
QY 408 TATATGGCTCGCAACCCCAAGATCTGCTGTGCTTCTTATTCAGTCCAGAGCTCCAAGTCA 467
DB 127 TyrLeuCysArgAsnAlaLysAspValAlaValSerTyrTyrPheLeuLeuMetIle 146
QY 468 CGGACCATGAGCTACCGGAGCACCTTTCAAGAAATTTCTCGCGGAGTTTATGAATGAAG 527
DB 147 LysSerTyrGlnAsnProLysSerPheSerGluPheValGluLysPheMetGluGlyGln 166
QY 528 CTGGCTACGGCTCTCTGTTTGTAGCACCTGAGGAGTTTCTGGGACACCGCATGGAAGT 587
DB 167 ValProTyrGlySerTrpTyrAspHisValLysSerTyrTrpGluLysSerLysAsnSer 186
QY 588 AACGTCCTTTTCTCAAGTATGAAGACATCGCATCGGACCTCGTGCAGTGTGGAGCAG 647
DB 187 ArgValLeuPheMetPheTyrGluAspMetLysGluAspIleArgArgGluValValLys 206
QY 648 CTGGCCAGATTCCTGGGGGTGTCTGTGACAAAGGCCAGCTCGAAGCCCTGAGGAGCAG 707
DB 207 LeuIleGluPheLeuGluArgAspProSerAlaGluLeuValAspArgIleIleGlnHis 226
QY 708 TGC-----CACGACCTGGTGGACCGAGCTGTGCTCAAGGCCAGCTCGAAGCCCTGAGGAGCAG 752
DB 227 ThrSerPheGlnGluMetLysAsnAsnProCysIleAsnTyrSerMetLeuProGluThr 246
QY 753 -----GTGGCCCGGGGAAGAGTTGGGCTGTGGAGGAGCAG 785
DB 247 MetIleAspLeuLysValSerProPheMetArgLysGlyIleValGlyAspTrpLysAsn 266
QY 786 ATCTTCACCGCTCCATGAATGAGAAGTTTTCAGTTGGTGTATTAACAACAAGATGGGAAG 845
DB 267 HisPheProGluAlaLeuArgGluArgPheGluHisTyrGlnGlnMetLysAsp 286
QY 846 TGTGACCTCAGCTTT 860
DB 846 TGTGACCTCAGCTTT 860
```


A:Molecule type: protein
A:Residues: 23-36;38-47;79-96;134-144;207-226;234-248 <SCW>
A:CROSS-References: UNIPARC:UPI0000175829; UNIPARC:UPI000017582A; UNIPARC:UPI000017582B;
C:Comment: This enzyme belongs to a family of cytosolic enzymes that catalyze the trans-
C:Genetics:
A:Gene: psc
A:Introns: 50/1; 93/3; 124/3; 167/1; 198/3; 259/1
C:Keywords: sulfotransferase

Alignment Scores:
Pred. No.: 3 91e-23 Length: 294
Score: 412.00 Matches: 91
Percent Similarity: 53.5% Conservative: 56
Best Local Similarity: 33.1% Mismatches: 108
Query Match: 9.4% Indels: 20
DB: 2 Gaps: 4

US-10-768-158-1 (1-2419) x JCS000 (1-294)

```
QY 69 AGCAAGTACTTCGAGTTCATCCATGCGGTGCGGTGCGCCCTTCTGCGCGGGAAGATGGAG 128
   :::::::::: :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
Db 12 AlaLysTyrrVallys-----GlylleProLeulleLysTyrrPheAlaGluAlaLeuGly 29
   :::::::::: :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
QY 129 GAGATGCCCAACTTCCGGTGGCGGCCAGGAGCTGTGGATCGTCACTACCCCAAGTCC 188
   :::::::::: :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
Db 30 ProLeuGluSerPheGluAlaTrpProAspAspLeuLeulleSerThrTyrrProLysSer 49
   :::::::::: :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
QY 189 GGCACCAAGTCTGCTGAGGAGGTGTCTACTTGTGTGAGCGCGCTCACCCCGATGAG 248
   :::::::::: :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
Db 50 GlyThrThrTrpValSerGluilleLeuAspLeulleTyrrGlnGluLysAspLeuGluLys 69
   :::::::::: :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
QY 249 ATCGGTTGATGAACATCAGCAGAGCTCCGGTCTCGAGTACCACAGCCG----- 302
   :::::::::: :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
Db 70 CysGlnArgAlaProValPheLeuArgValProPheLeuGluPheSerAlaProGlyVal 89
   :::::::::: :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
QY 303 -----GGCTGACATCATCAAGAACTGACCTCTCCCGCTCATCAAGACCCACCTG 356
   :::::::::: :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
Db 90 ProThrGlyValGluLeulleLysAspThrProAlaProArgLeulleLysThrHisLeu 109
   :::::::::: :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
QY 357 CCTTACCGCTTCTGCGCTTCCACCTCCCAATGGAGACTCCAAAGGTCTATATATGGCT 416
   :::::::::: :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
Db 110 ProLeuAlaLeuLeuProLysThrLeuLeuAspGlnLysVallyleTyrrleAla 129
   :::::::::: :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
QY 417 CGCAACCCCAAGATCTGTGTGTCTTATATCACTTCACTGACCGCTCTCTGCGGACCATG 476
   :::::::::: :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
Db 130 ArgAsnAlaLysAspValAlaValSerTyrrHisPheTyrrArgMetAlaLysValHis 149
   :::::::::: :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
QY 477 AGCTACCGAGGACCTTCAAGAACTTCCGGAGGTCTGAGGAGTTTATGAATGATGAGTGGCTAC 536
   :::::::::: :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
Db 150 ProAspProGlyThrTrpAspSerPheLeuGluLysPheMetAlaGlyGluValCysTyrr 169
   :::::::::: :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
QY 537 GGTCTCTGTTTGGACACGTCGAGGAGTCTGGGAGCACCGCATGGACTCGAAGCTGCTT 596
   :::::::::: :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
Db 170 GlySerTrpTrpGlnHisValGlnGluTrpTrpGluLeuSerHisThrHisProValLeu 189
   :::::::::: :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
QY 597 TTCTCAAGTATGAACATGCACTCGGACCTCGTGCAGCATGGTGGAGCAGCTGGCCACA 656
   :::::::::: :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
Db 190 TyrLeuPheTyrrGluAspLysGluAspProLysArgGluilleGlnLysleLeuGlu 209
   :::::::::: :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
QY 657 TTCTCGGGGTCTCTGTGACAGGCCAGCTGCGGAGCCCTG-----ACGGAGCAC 707
   :::::::::: :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
Db 210 PheilleGlyArgSerLeuProGluGluThrValAspHisleValGlnArgThrSerPhe 229
   :::::::::: :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
QY 708 TGCCACCAAGTGTGTCAGACAGTCTGCAACGCTGAGGCCCTGCGCGTGGGC----- 758
   :::::::::: :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
Db 230 LysGluMetLysLysAsnProMetThrAsnTyrrSerThrleProThrAlaValMetAsp 249
   :::::::::: :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
QY 759 -----CGGGGAAGAGTTGGGTGTGGAGGACATCTTCACC 794
   :::::::::: :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
Db 250 HisSerileSerAlaPheMetArgLysGlyleThrGlyAspTrpLysSerThrPheThr 269
   :::::::::: :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
QY 795 GTCTCCATGAATCAGAGTTTGATTTGGTGTATATAACAGAGATG 839
```

Db 270 ValAlaGlnAsnGluLeuPheGluAlaHisTyrrAlaLysLysMet 284
RESULT 27
A44011
adrenocortical estrogen sulfotransferase - guinea pig
C:Species: Cavia porcellus (guinea pig)
C:Date: 27-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: A44011
R:Oeda, T.; Lee, Y.C.; Driscoll, W.J.; Chen, H.C.; Strott, C.A.
Mol. Endocrinol. 6, 1216-1226, 1992
A:Title: Molecular cloning and expression of a full-length complementary DNA encoding the
A:Reference number: A44011; MUID:93024479; PMID:1406700
A:Accession: A44011
A:Status: preliminary
A:Molecule type: mRNA; protein
A:Residues: 1-296 <OED>
A:CROSS-References: UNIPROT:P49887; UNIPARC:UPI0000136224; GB:U09552; GB:S45979; NID:9491
A:Note: sequence extracted from NCBI backbone (NCBIN:115159, NCBIP:115160)
C:Superfamily: alcohol sulfotransferase

Alignment Scores:
Pred. No.: 7.23e-23 Length: 296
Score: 408.50 Matches: 94
Percent Similarity: 52.4% Conservative: 50
Best Local Similarity: 34.2% Mismatches: 112
Query Match: 9.3% Indels: 19
DB: 2 Gaps: 4

US-10-768-158-1 (1-2419) x A44011 (1-296)

```
QY 72 AGTACTTCTC---GAGTTCATGCGGTGCGGTGCGCCCTTCTGCGCGGGAAGATGGAG 128
   :::::::::: :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
Db 11 GluTyrrPheAspGluPheArgGlylleLeulleTyrrLysGlnPheLysTyrrTrpAsp 30
   :::::::::: :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
QY 129 GAGATGCCCAACTTCCGGTGGCGGCCAGCAGCTGTGGATCGTCACTTACCCCAAGTCC 188
   :::::::::: :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
Db 31 AsnValGluAlaPheGlnAlaArgProAspAspLeuValleAlaAlaTyrrProLysSer 50
   :::::::::: :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
QY 189 GGCACCAAGTCTGCTGAGGAGTGTCTACTTGTGTGAGCGCGCTGACCCCATGAG 248
   :::::::::: :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
Db 51 GlyThrThrTrpIleSerGluValValCysMetIleTyrrAlaGluLysVallyLys 70
   :::::::::: :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
QY 249 ATCGGCTTCATCAACATCCACAGCAGCTCCCGGTCTCGGACTCCACAGCCG----- 302
   :::::::::: :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
Db 71 CysArgGlnAspAlaIlePheAsnArgValProPheLeuGluCysArgAsnAspLysMet 90
   :::::::::: :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
QY 303 -----GGCTGACATCATCAAGAACTGACCTCTCCCGCTCATCAAGACCCACCTG 356
   :::::::::: :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
Db 91 MetAsnGlyValLysGlnLeuGluGluMetAsnSerProArgIlelleLysThrHisLeu 110
   :::::::::: :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
QY 357 CCTTACCGCTTCTGCGCTTCCACCTCCCAATGGAGACTCCAAAGTCTATATATGGCT 416
   :::::::::: :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
Db 111 ProProArgLeuLeuProAlaSerPheTrpGluLysArgCysLysMetIleCysleCys 130
   :::::::::: :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
QY 417 CGCAACCCCAAGATCTGTGTGTCTTATATCACTTCCACCGCTCTCTGCGGACCATG 476
   :::::::::: :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
Db 131 ArgAsnAlaLysAspValAlaValSerTyrrTyrrPhePheLeuMetValAlaAsnHis 150
   :::::::::: :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
QY 477 AGCTACCGAGGACCTTTCAGAAATTCGCCGAGGTTTATGAATGATGAGTGGGCTAC 536
   :::::::::: :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
Db 151 ProAspProGlySerPheProGluPheValGluLysPheMetGlnGlyGlnValProTyrr 170
   :::::::::: :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
QY 537 GGTCTCTGTTTGGACACGTCGAGGAGTCTCGGAGCAGCCGATGAGTCAAGACGTGCTT 596
   :::::::::: :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
Db 171 GlySerTrpTyrrAspHisVallySerTrpTrpGluLysSerThrAspProArgIleLeu 190
   :::::::::: :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
QY 597 TTCTCAAGTATGAACATGCACTCGGACCTGCGGAGCCCTG-----ACGGAGCAC 656
   :::::::::: :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
Db 191 PheillePheTyrrGluAspMetLysGluAspIleArgLysGluValLeuLysleulleHis 210
   :::::::::: :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
QY 657 TTCTCGGGGTGTCTGTGACAGGCCAGCTGCGGAGCCCTGAGGAGCACTGCCAC--- 713
   :::::::::: :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
```

Db 211 PheLeuGlyArgLysProSerGluGluLeuValAspLysIleIleLysHisThrSerPhe 230
 QY 714 -----CAGCTGGTGACCAAGTGTGCAAC 737
 Db 231 GlnGluMetLysAsnAsnProSerThrAsnTyrThrMetLeuProGluIleMetAsn 250
 QY 738 GCTGAGCCCTGCC--GTGGCCGGGAGAGTTGGGCTGTGGAGGACATCTTCACC 794
 Db 251 GlnLysValSerProPheMetArgLysGlyIleSerGlyAspTrpLysAsnHisPheThr 270
 QY 795 GTCTCCATGAATGAGAAGTTTGACTTGGTGTATAAACAGAGATG 839
 Db 271 ValAlaLeuAsnGlnSerPheAspLysHisTyrGlnGlnMet 285
 RESULT 28
 JW0078
 amine sulfotransferase (EC 2.8.2.3) RB1 - rabbit
 N:Alternate names: AST-RB1
 C:Species: Oryctolagus cuniculus (domestic rabbit)
 C>Date: 17-Jun-1998 #sequence_revision 10-Jul-1998 #text_change 09-Jul-2004
 C:Accession: JW0078
 R:Yoshinari, K.; Nagata, K.; Ogino, M.; Fujita, K.; Shiraga, T.; Iwasaki, K.; Hata, T.; J. Biochem. 123, 479-486, 1998
 A:Title: Molecular cloning and expression of an amine sulfotransferase cDNA: a new gene
 A:Reference number: JW0078, MUID:98207028; PMID:9538231
 A:Accession: JW0078
 A:Molecule type: mRNA
 A:Residues: 1-301 <YOS>
 A:Cross-references: UNIPROT:O46640; UNIPARC:UPI0000086C35; DDBJ:D86219; NID:g2916982; PI
 A:Experimental source: liver
 C:Comment: This protein catalyzes 4-phenyl-1,2,3,6-tetrahydropyridine sulfation.
 C:Superfamily: alcohol sulfotransferase
 C:Keywords: sulfotransferase

Alignment Scores:
 Pred. No.: 4,97e-22 Length: 301
 Score: 397.50 Matches: 85
 Percent Similarity: 55.4% Conservative: 68
 Best Local Similarity: 30.8% Mismatches: 105
 Query Match: 9.0% Indels: 18
 DB: 2 Gaps: 4

US-10-768-158-1 (1-2419) x JW0078 (1-301)

QY 123 ATGGAGGAGATCCCAACTTCCTCCGTGCGGCCGACGAGCGTGTGGATCTCACTACCCC 182
 Db 26 LeuGluLysLeuAspAspPheGluIleArgAspAspValPheValIleThrTyrPro 45
 QY 183 AGTCCGGCACCAGCTTGCTGTCAGGAGGTGTCTACTTGGTGACCGCGCTGACCCC 242
 Db 46 LysSerGlyThrValTrpThrGlnGlnIleLeuSerLeuIleTyrPheGluGlyHisArg 65
 QY 243 GATGAGATCGGCTTGATGAACATCGACGAGCAGCTCCCGTCTCTGGAGTACCCACAGCG 302
 Db 66 AsnArgThrGluLysTrpAspThrLeuAspArgValProPheLeuGluTyrAsnIleArg 85
 QY 303 GGCCTGGACATCATCAAGAACTACCTTCCTCCCGCTCATCAAGAGCCACCTGCGCTAC 362
 Db 86 LysValAsp--IleGluAsnArgProSerProArgLeuPheAlaSerHisLeuProTyr 104
 QY 363 CGCTTTCCTCTGACCTCCACATGGAGACTCCAGGTCTATATATGCTGCTCAAC 422
 Db 105 TyrLeuAlaProLysSerLeuLysAsnAsnLysAlaLysIleIleTyrValTyrArgAsn 124
 QY 423 CCCAAGGATCTGGTGTCTTATTATATGATTTCCACCGCTCTCTCGGACCATGAGCTAC 482
 Db 125 ProLysAspValIleIleSerPheHisPheSerAsnMetValValLysLeuGluAla 144
 QY 483 CGAGGCACCTTTCAAGAATCTCGCGAGGTTTATGAATGATAAGTGGGCTACGGCTCC 542
 Db 145 SerAsnThrLeuGluAsnPheMetGluLysPheLeuAspGlyLysValValGlySerIle 164
 QY 543 TGGTTTGACACGTCAGGAGTTCTGGGAGCACCGCATGGATCGAACCTGCTTTTCTC 602

Db 165 TrpPheAspHisIleArgGlyTrpTyrGluHisLysAsnAspPheAsnIleLeuPheMet 184
 QY 603 AAGTATGAAGACATCATCGGACCTGTGTGACGATGTGGAGCAGCTGCCACAGATTCCCTG 662
 Db 185 MetTyrGluAspMetLysLysAspLeuArgSerSerIleLeuLysIleSerPheLeu 204
 QY 663 GGGGTGTCTGTGACAAAGCCCGACCTGGAGCCCTGACGGACACTGC----- 710
 Db 205 GluLysAspLeuSerGluGluValAspAlaIleValArgGlnAlaThrPheGluAsn 224
 QY 711 CACCAGCTGGTGACACAGTGC-----TGCACGCTGAGGCCCTGCCCTGGGCCGG--- 761
 Db 225 MetLysPheIleProGlnAlaAsnTyrAsnAsnIleLeuSerAsnGluIleGlyArgArg 244
 QY 762 -----GGAAGAGTTGGCTGTGTGGAAGGACACTTTCACC 794
 Db 245 HisAsnGluGlyAlaPheLeuArgLysGlyAlaValGlyAspTrpLysHisMetThr 264
 QY 795 GTCTCCATGAATGAGAAGTTTGACTTGTGTATTAACAGACAGATGGGAAGTGTGACCTC 854
 Db 265 ValGluGlnSerGluArgPheAspArgIlePheGlnGluGluMetLysAspPheProLeu 284
 QY 855 ACCTTTTGACTTTTATTATTAATAACAGAAACAAACCACTGCATGCT 900
 Db 285 LysPheIleTrpAspLeu-AsnAspGluAlaAsnSerAsnHisSer 299
 RESULT 29
 JC4531
 alcohol sulfotransferase (EC 2.8.2.2) 2 - guinea pig
 N:Alternate names: adrenal hydroxysteroid sulfotransferase 2
 C:Species: Cavia porcellus (guinea pig)
 C>Date: 15-Feb-1996 #sequence_revision 19-Apr-1996 #text_change 09-Jul-2004
 C:Accession: JC4531; PC4120
 R:Lu, N.X.; Driscoll, W.J.; Martin, B.M.; Strott, C.A. Biochem. Biophys. Res. Commun. 217, 1078-1086, 1995
 A:Title: Molecular cloning and expression of a guinea pig 3-hydroxysteroid sulfotransferase
 A:Reference number: JC4531; MUID:96125350; PMID:8554560
 A:Accession: JC4531
 A:Molecule type: mRNA
 A:Residues: 1-287 <LUU>
 A:Cross-references: UNIPROT:P52841; UNIPARC:UPI0000170746; GB:U35115; NID:g1151080; PID:M
 A:Experimental source: adrenal gland
 A:Accession: PC4120
 A:Molecule type: protein
 A:Residues: 65-73;75-83;95-103;206-215;274-287 <LU2>
 A:Cross-references: UNIPARC:UPI000017581C; UNIPARC:UPI000017581D; UNIPARC:UPI000017581E,
 C:Comment: This enzyme sulfonates both 3 alpha- and 3 beta-hydroxylated neutral steroids
 C:Genetics:
 A:Gene: gpHST2
 A:Superfamily: alcohol sulfotransferase
 C:Keywords: adrenal gland; sulfotransferase
 P:249-255/Region: DNA binding #status predicted

Alignment Scores:
 Pred. No.: 4.01e-21 Length: 287
 Score: 385.50 Matches: 88
 Percent Similarity: 53.4% Conservative: 55
 Best Local Similarity: 32.8% Mismatches: 108
 Query Match: 8.8% Indels: 17
 DB: 2 Gaps: 5

US-10-768-158-1 (1-2419) x JC4531 (1-287)

QY 84 TTCCATGCGTGGCTGCTCCG-----CCCTTCTCCCGGGGAAGATGGAGAGATCGCC 137
 Db 8 PheGluGlyIleArgPheProMetValGlyPheSerProGluLeuLeuArgGluValArg 27
 QY 138 AAC---TTCCCGGTGCGGCCCGCAGCGAGTGTGGATCGTCACTACCCCAAGTCCGGCACC 194
 Db 28 AspLysPheLeuValLysAspGluAspThrIleThrValThrTyrProLysSerGlyThr 47
 QY 195 ACCTTGCTGCAGGAGGTGCTACTTGTGTAGCCAGCGCGCTGACCCCGCATGAGATCGGC 254

A:Accession: A54026

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-287 <LEE>

A:Cross-references: UNIPROT:P50234; UNIPARC:UPI0000170743; GB:U06871; NID:g508212; PIDN:

C:Superfamily: alcohol sulfotransferase

C:Keywords: sulfotransferase

Alignment Scores:
Pred. No.: 1.93e-20 Length: 287
Score: 376.50 Matches: 88
Percent Similarity: 51.1% Conservative: 49
Best Local Similarity: 32.8% Mismatches: 114
Query Match: 17 Indels: 17
DB: 2 Gaps: 5

US-10-768-158-1 (1-2419) x A54026 (1-287)

```
QY 84 TTCATGGCGTGGCGTGGCGCCCTTCGCCGGGAAG-----ATGAGGAGATCGCC 137
Db 8 PheGluGlyIleArgPheProMetValCysPheSerProGluIleuArgGluValArg 27

QY 138 AAC---TTCGCCGTGGCGCCAGCAGCGTGTGGATCGTCACCTACCCCAAGTCCGCGACC 194
Db 28 AspLysPheLeuValuAspGluAspThrIleThrValThrTyProLysSerGlyThr 47

QY 195 AGCTTGCTGCAGAGGTGCTCTACTTGGTGAGCCAGCGCGCTGACCCCGATGAGATCGGC 254
Db 48 AsnTrpLeuAsnGluIleValCysLeuIleLeuSerLysGlyAspProLysLeuValGln 67

QY 255 TTGATGAACATCGACGAGCAGCTCCGGTCTGGAGTACCCACAGCCGGCCCTGGACATC 314
Db 68 SerValProAsnTrpAspArgSerProTrpIleGluPhe---ThrGlyGlyTyrgluLeu 86

QY 315 ATCAAGGAACCTGACCTCTCCCGCTCATCAAGAGCCAGCTCCCTACCGCTTCTGCCC 374
Db 87 ValLysGlyGlnLysAspProArgValTyThrSerHisLeuProLeuHisLeuPhePro 106

QY 375 TCTGACCTCCAAATGGAGACTCCAAGGTCACTATATGGTTCGCAACCCCAAGGATCTG 434
Db 107 LysSerPhePheSerSerLysAlaLysValIleTyrcysIleArgAsnProArgAspAla 126

QY 435 GTGGTGCTTATTACGTTCACCGCTCTCTGCGGACCATGAGTACCGAGGACACCTTT 494
Db 127 LeuValSerGlyTyPhePheLeuSerLysMetAsnValThrGluLysProGluThrLeu 146

QY 495 CAAGAATTCTGCGGAGGTTTATGAATGATAAGCTGGGTACGGCTCCTGTTTTCAGCAC 554
Db 147 GlnGlnTyMetGluTrpPheLeuGlnGlyAsnValIleTyrglySerTrpPheGluHis 166

QY 555 GTCCAGGAGTTCTGGGAGCACCGCATGGACTCGAACGTGCTTTTCTCAAGTATGAAGAC 614
Db 167 ValArgGlyTyTrpLeuSerMetArgGluMetGluAsnValLeuValSerTyrgluAsp 186

QY 615 ATGCATCGGACCTGGTACGATGTGGAGCAGCTGGCCAGATTCTGGGGGTGTCCTGT 674
Db 187 LeuIleLysAspThrArgSerThrValGluLysIleCysGlnPheLeuGlyLysLysLeu 206

QY 675 GACAGGCCAGCTGGAACCTGACGAGCACTGC-----CACCAGCTGTGTGAC 725
Db 207 LysProGluGluThrAspLeuValLeuLysTySerSerPheGlnPheMetLysGluAsn 226

QY 726 CAGTGCTGCAACGCTGAGCGCCCTGCC-----GTG 755
Db 227 GluMetSerAsnPheThrLeuLeuProHisAlaTyThrThrGluGlyPheThrLeuLeu 246

QY 756 GCCCGGGAAGATTGGGTGGGAAGACATCTTTCACCGTCTCCATCAATGAGAAGTTT 815
Db 247 ArgLysGlyThrValGlyAspTrpLysAsnHisPheThrValAlaGlnAlaGluAlaPhe 266

QY 816 GACTTGGTGATAAACAGAGATG 839
Db 267 AspLysIleTyrglnGluLysMet 274
```

RESULT 32

JE0152

alcohol sulfotransferase (EC 2.8.2.2) - rabbit

N/Alternate names: ASF-RB2; hydroxysteroid sulfotransferase

C:Species: Oryctolagus cuniculus (domestic rabbit)

C:date: 13-Jun-1998 #sequence_revision 10-Jul-1998 #text_change 09-Jul-2004

C:Accession: JE0152

R:Yoshinari, K.; Nagata, K.; Shiraga, T.; Iwasaki, K.; Hata, T.; Ogino, M.; Ueda, R.; F.

J. Biochem. 123, 740-746, 1998

A:Title: Molecular cloning, expression, and enzymatic characterization of rabbit hydroxy-

A:Reference number: JE0152; MUID:98207066; PMID:9538269

A:Accession: JE0152

A:Molecule type: mRNA

A:Residues: 1-286 <YOS>

A:Cross-references: UNIPROT:O62648; UNIPARC:UPI00000886AB; DDBJ:AB006053; NID:g3036938;

A:Experimental source: liver

C:Comment: This protein shows high activities to both hydroxysteroids and amines.

C:Superfamily: alcohol sulfotransferase

C:Keywords: sulfotransferase

Alignment Scores:

Pred. No.: 5.06e-20 Length: 286
Score: 371.00 Matches: 90
Percent Similarity: 51.5% Conservative: 51
Best Local Similarity: 32.8% Mismatches: 113
Query Match: 8.4% Indels: 20
DB: 2 Gaps: 6

US-10-768-158-1 (1-2419) x JE0152 (1-286)

```
QY 63 TTCGAGACCAAGTACTTCAGTTCATGGCGTGGCGCTTCTTCCGCGGGGAAG 122
Db 8 TyrGluGlyIleAlaPheProPheValGlyPheAsn---ProGluLeuMetArgLysAla 26

QY 123 ATGAGGAGATCGCAACTTCCCGTGGCGGCCAGCAGCTGTGTGATCGTACCTACCCC 182
Db 27 TyrGluGlu-----PheValValAsnGluAspValLeuThrValThrPhePro 43

QY 183 AAGTCGGCCACAGCTTCTGTCAGGAGGTGTCTACTTTGGTGAGCCAGCGCTGACCCC 242
Db 44 LysSerGlyThrAsnTrpLeuIleGluLeuLeuCysLeuIleArgSerLysGlyAspAla 63

QY 243 GATGAGATCGGCTTGATGAACATCGACGAGCAGCTCCCGCTCCTGGAGTACCCACAGCCG 302
Db 64 ThrGlnIleGlnSerValProIleTrpValArgSerProTrpValGlu---ThrValSer 82

QY 303 GGCCTGGACATCATCAAGAACTGACCTCTCCCGCTCATCAAGAGCCACCTGCGCTAC 362
Db 83 GlyTyrGluAspLeuLysThrMetGluSerProArgLeuIleSerThrHisLeuProIle 102

QY 363 CGCTTTCTGCCCTCTGACCTCCACAATGGAGACTCCAAAGGTCACTATATGGCTGCAAC 422
Db 103 HisIlePheProLysSerIleHisThrThrLysAlaLysValIleTyLeuMetArgAsn 122

QY 423 CCCAAGGATCTGGTGTGCTTATTATCAGTTTCCACCGCTCTCTGCGGACCATGAGTAC 482
Db 123 ProArgAspValLeuValSerGlyTyTrpPheTrpAsnTyTrpValLysPheValGluAsn 142

QY 483 CGAGGCACCTTTCAGAAATTCGCGGAGGTTTATGAATGATAGCTGGGTACGGCTCC 542
Db 143 ThrLysSerLeuGlnGluTyPheGluLeuPheLeuAspGlyAsnValValPheGlySer 162

QY 543 TGGTTTGAGCAGCTGAGGAGTTCCTGGGAGCACCGCATGGACTCGAAGCTCTTTTCTC 602
Db 163 TrpPheAspHisValHisGlyTrpLeuSerLeuArgAspGlnLysAsnPheLeuLeu 182

QY 603 AAGTATGAACATGATCATCGGACCTGGTGTGACGATGGTGGAGCAGCTGGCCAGATTCT 662
Db 183 SerTyrgluGluLeuLysGlnAspThrArgSerThrIleGluLysIleCysHisPheLeu 202

QY 663 GGGGTGTCTGTGACAAAGGCCACGCTG-----GAAGCC 695
Db 663 GGGGTGTCTGTGACAAAGGCCACGCTG-----GAAGCC 695
```


Db	28	AspGluPheValIleAargAspGluAaspValIleLeuThrTyrProLysSerGlyThr	47
QY	195	AGCTTGTCTCAGAGAGTGGTCTACTTTGGTGAGCCAGGGCGCTCACCCCGATGAGATCGGC	254
Db	48	AsnTrpLeuAlaGluIleLeuCysLeuMetHisSerLysGlyAspAlaLysTrpIleGln	67
QY	255	TTGATGAACATCAGCAGCAGCTCCGGTCTGGAGTACCCACAGCCGGCCCTGGACATC	314
Db	68	SerValProIleTrpGluAargSerProTrpValGlu---SerGluIleGlyTyrThrAla	86
QY	315	ATCAAGGAAGTCACTCTCCCGCTCATCAAGAGCCACTGCCTTACCGCTTCTTGCC	374
Db	87	LeuSerGluThrGluSerProAargLeuPheSerSerHisLeuProIleGlnLeuPhePro	106
QY	375	TCTGACTCCCAATGAGACTCCAAAGTCACTATATATGGCTCGCAACCCCAAGGATCG	434
Db	107	LysSerPhePheSerSerLysAlaLysValIleTyrLeuMetArgAsnProArgAspVal	126
QY	435	GTGGTGTCTTATATCAGTTCCACCGCTCTCTGGCGGACCATGACTACCGAGCACTTT	494
Db	127	LeuValSerGlyTyrPhePheTrpLysAsnMetLysPheIleLysLysProLysSerTrp	146
QY	495	CAAGAATCTGCGGAGGTTTATGAATGATAAGCTGGGCTACGGCTCCTGGTTTGAGCAC	554
Db	147	GluGluTyrPheGluTrpPheCysGlnGlyThrValLeuTyrGlySerTrpPheAspHis	166
QY	555	GTGAGGAGTTCGGGAGCACCGCATGGAAGTCTGCTTTCTTCAAGTATGAAGAC	614
Db	167	IleHisGlyTrpMetProMetArgGluGluLysAsnPheLeuLeuLeuSerTyrGluGlu	186
QY	615	ATGCATCGGACCTGGTACGATGGTGAGCAGCTGGCCAGATTTCTGGGGGTGTCGTGT	674
Db	187	LeuLysGlnAspThrGlyArgThrIleGluLysIleCysGlnPheLeuGlyLysThrLeu	206
QY	675	GACAAGGCCACGTGGAAGCCCTGACGGAGCACTGC-----CACCAGCTGTGGAC	725
Db	207	GluProGluGluLeuAsnLeuIleLysAsnSerSerPheGlnSerMetLysGluAsn	226
QY	726	CAGTCTGCAACGCTGAGGCCCTGCCGTGGGC-----	758
Db	227	LysMetSerAsnTyrSerLeuLeuSerValAspTyrValValAspLysAlaGlnLeuLeu	246
QY	759	---CGGGAAGATTGGGCTGTGGAGGACATCTTCACCGTCTCCATGATGAGAAGTTT	815
Db	247	ArgLysGlyValSerGlyAspTrpLysAsnHisPheThrValAlaGlnAlaGluAspPhe	266
QY	816	GACTTGGTGATAAACAGAGATG	839
Db	267	AspLysLeuPheGlnGluLysMet	274
RESULT 37			
152849			
alcohol sulfotransferase (EC 2.8.2.2) - rat			
N:Alternate names: hydroxysteroid sulfotransferase			
C:Species: Rattus norvegicus (Norway rat)			
C>Date: 12-Aug-1996 #sequence revision 12-Aug-1996 #text_change 09-Jul-2004			
C:Accession: 152849; 165759			
R:Watabe, T.; Ogura, K.; Satsukawa, M.; Okuda, H.; Hiratsuka, A.			
Chem. Biol. Interact. 92, 87-105, 1994			
A>Title: Molecular cloning and functions of rat liver hydroxysteroid sulfotransferases c			
A:Reference number: 152849; MUID:94306585; PMID:8033273			
A:Accession: 152849			
A>Status: preliminary; translated from GB/EMBL/DBJ			
A:Molecule type: mRNA			
A:Residues: 1-284 <RES>			
A:Cross-references: UNIPROT:P15709; UNIPARC:UPI0000170BD9; GB:D14987; NID:G303797; PIDN:			
C:Superfamily: alcohol sulfotransferase			
C:Keywords: sulfotransferase			
Alignment Scores:			
Pred. No.:	1,08e-18	Length:	284
Score:	353.50	Matches:	89
Percent Similarity:	52.2%	Conservative:	54

Best Local Similarity: 32.5%			
Query Match: 8.0%			
DB:	2	Mismatches:	108
		Indels:	23
		Gaps:	7
US-10-768-158-1 (1-2419) x 152849 (1-284)			
QY	75	TACTTCGAGTTCATCGGCTGCGGCTGCGGCCCTTCTGCGCGGGAAG-----ATGGAG	128
Db	4	TyrThrTrpPheGluGlyIleProPheHisAlaPheGlyIleSerLysGluThrLeuGln	23
QY	129	GAGATCGCCAAC---TTCCTGGTGGCGCCAGCAGCTGTGGATCGTCACTACCCACAG	185
Db	24	AenValCysAsnLysPheValValLysGluGluAspLeuLeuAlaTyrProLys	43
QY	186	TCCGCAACCAAGTCTCTGAGGAGTGGTCTACTTTGGTGAGCCAGGGCGCTCACCCCGAT	245
Db	44	SerGlyThrAsnTrpLeuIleGluIleValCysLeuIleGlnThrLysGlyAspProLys	63
QY	246	GAGATCGGCTTGATGAACATGACGAGCAGCTCCCGTCTCGGATGATGATGATGATGATG	305
Db	64	TrpIleGlnSerValThrIleTrpAspArgSerProTrpIleGlu---ThrAspValGly	82
QY	306	CTGGACATCAACGAAGTCACTCTCCCGCTCATCAAGAGCCAGCCCTGCGCTACCCG	365
Db	83	TyrAspIleLeuLeuLysLysGlyProArgLeuMetThrSerHisLeuProMetHis	102
QY	366	TTTCTGCGCTCTGACTCCCAATGGAGACTCCAAAGTCACTATATGCTCGCAACCCC	425
Db	103	LeuPheSerLysSerLeuPheSerSerLysAlaLysValIleTyrLeuIleArgAsnPro	122
QY	426	AAGATCTGGTGGTCTTATATCAGTTCCACCCCTCT-----CTGCGGACCATG	476
Db	123	ArgAspValLeuValSerGlyTyrTyrPheTrpGlyAsnSerThrLeuValLysPro	142
QY	477	AGCTACCCAGCAGCTTTCAGAAATTCGCGGAGGTTTATGAATGAATAGCTGGGCTAC	536
Db	143	AspSerLeuGlyThrTyrValGluTrp-----PheLeuLysGlyAsnValLeuTyr	159
QY	537	GGCTCTCTGTTGACGAGTCTTCTGGAGCACCGCATGAGTCTGCAACGCTGCTT	596
Db	160	GlySerTrpPheGluHisIleArgAlaTrpLeuSerMetArgGluTrpAspAsnPheLeu	179
QY	597	TTTCTCAAGTATGAAGATCATCGGACCTGTGTGACGATGTTGGAGCAGCTGGCCAGA	656
Db	180	LeuLeuTyrTyrGluAspMetLysLysAspThrMetGlyThrIleLysLysIleCysAsp	199
QY	657	TTCTTGGGGTGTCTGTGACAAAGCCAGCTGGAGCCCTGACGAGCAGCTGCCACCAG	716
Db	200	PheLeuGlyLysLysLeuGluProAspGluLeuAspLeuValLeuLysTyrSerSerPhe	219
QY	717	CTGGTG-----GACCACTGTGCAACCTGAGGCCCTGCCCC-----	752
Db	220	GlnValMetLysGluAsnAspMetSerAsnTyrSerLeuLeuMetLysLysSerIlePhe	239
QY	753	-----GTGGCGCGGGAAGATTTGGCTGTGGGAAGACATCTTCAACGCTC	797
Db	240	ThrGlyIleGlyLeuMetArgLysGlyThrIleGlyAspTrpLysAsnHisPheThrVal	259
QY	798	TCCATGATGAGAAGTTTGCATTTGGTGTATATAACAGAGATG	839
Db	260	GlnAlaGluAlaPheAspLysValPheGlnGluLysMet	273
RESULT 38			
A33569			
alcohol sulfotransferase (EC 2.8.2.2) - rat			
C:Species: Rattus norvegicus (Norway rat)			
C>Date: 09-Mar-1990 #sequence_revision 09-Mar-1990 #text_change 09-Jul-2004			
C:Accession: A33569			
R:Ogura, K.; Kajita, J.; Narihata, H.; Watabe, T.; Ozawa, S.; Yamazoe, Y.;			
Biochem. Biophys. Res. Commun. 165, 168-174, 1989			
A>Title: Cloning and sequence analysis of a rat liver cDNA encoding hydroxysteroid sulfo			
A:Reference number: A33569; MUID:90073705; PMID:2590219			
A:Accession: A33569			

A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-284 <OG>
A;Cross-references: UniProt
C;Superfamily: alcohol dehydrogenase
C;Keywords: sulfotransferase

Alignment Scores:

Length:	1,536-18	284
Matches:	351.50	90
Conservative:	52.7%	55
Mismatches:	32.7%	105
Indels:	8.0%	25
Gaps:	2	8
NA:		

US-10-768-158-1 (1-2419) x A33569 (1-284)

QY	75	TACTTCGAGTTCCATGGGGTGGCGGTGCGCGCCCTTCTGCCGGGGAG-----ATGGAG	120
DB	4	TyrThrTrpPheGluGlyLeuProPheHisAlaPheGlyIleSerIysGluThrLeuGln	23
QY	129	GAGATCGCCAAC---TTCGGTGGCGGCCACGCGAGCTGTGATCGTACCTACCTACCCCAAG	185
DB	24	AsnValCysAsnLysPheValValLysAspGluAspLeuLeuLeuAlaTyrProLys	43
QY	186	TCCGGCACACAGCTGCTGCGAGAGGTGGTCTACTTTGGTGAGCGAGCGGCGCTGACCCCGAT	245
DB	44	SerGlyThrAsnTrpLeuLeuGluLeuValCysLeuLeuGlnThrLysGlyAspProLys	63
QY	246	GAGATCGGCTTGATGAACATCGACGAGCAGCTCCCGGTCTGGAGTACCCACAGCGCGGC	305
DB	64	TrpIleGlnSerValThrIleTrpAspArgSerProTrpIleGlu---ThrAspValGly	82
QY	306	CTGGACATCATCAAGGAACGTGACCTCTCCCGCTCATCAAGAGCCACCTGCGCTACCGC	365
DB	83	TyrAspIleLeuLeuLysLysLysGlyProArgLeuMetThrSerHisLeuProMetHis	102
QY	366	TTTCTGCCCTCGACCTCCACATGGAGACTCCAAGTCACTCATATGCTGCGAACCC	425
DB	103	LeuPheSerLysSerLeuPheSerLysAlaLysValIleTyrLeuValArgAsnPro	122
QY	426	AAGGATCTGGTGGTCTTATTATCATGTTCCACCGCTCTCTGCGGACCATGAGCTACCGA	485
DB	123	ArgAspValLeuValSerGlyTyrTyrPheTrpGlyAsn---SerThrLeuAlaLysLys	141
QY	486	-----GGCACCTTCAAGAAATCTGCGGAGGTTTATGAATGATAGCTGGCG	533
DB	142	ProAspSerLeuGlyThrTyrValGluTrp-----PheLeuLysGlyAsnValLeu	158
QY	534	TACGGCTCTGTTGAGCAGTGCAGGAGTTCCTGGGACACCGCATGAGCTCGACGTG	593
DB	159	TyrGlySerTrpPheGluHisIleArgAlaTrpLeuSerMetGlnGluTrpAspAsnPhe	178
QY	594	CTTTTCTCAAGTATCAAGACATGCATCGGACCTGTGTGACGATGTGGAGCAGCTGGCC	653
DB	179	LeuLeuLeuTyrTyrGluAspMetLysLysAspThrMetGlyThrIleLysLysIleCys	198
QY	654	AGATTCTCGGGGTGCTCTGTGACAAAGCCCGACCTGGAGCCCTGACGGAGCACTGCCAC	713
DB	199	AspPheLeuGlyLysLysLeuGluProAspGluLeuAspLeuValLeuLysTyrSerSer	218
QY	714	CAGCTGGTG-----GACCAGTGTGCAACGCTGAGCGCCCTGCGCC	752
DB	219	PheGlnValMetLysGluAsnAspMetSerAsnTyrSerLeuLeuMetLysLysSerIle	238
QY	753	-----GTGGCCCGGGGAAGTGGGCTGTGGGAAGGACATCTTCACC	794
DB	239	PheThrGlyThrGlyLeuMetArgLysGlyThrValGlyAspTrpLysAsnHisPheThr	258
QY	795	GTCTCCATGAATGAGAAGTTTGACTTTGGTGTATATAACAGAAAGATG	839
DB	259	ValSerGluAlaGluAlaPheAspLysValPheGlnGluLysMet	273

RESULT 39

A26136
 sensequence marker protein 2, hepatic - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 05-Oct-1988 #sequence_revision 05-Oct-1988 #text_change 09-Jul-2004
 C:Accession: A26136
 E:Chatterjee, B.; Majumdar, D.; Ozbilen, O.; Murty, C.V.R.; Roy, A.K.
 J. Biol. Chem. 262, 822-825, 1987
 A:Title: Molecular cloning and characterization of cDNA for androgen-repressible rat liv
 A:Reference number: A26136; MUID:87109186; PMID:3805009
 A:Accession: A26136
 A:Molecule type: mRNA
 A:Residues: 1-282 <CHA>
 A:Cross-references: UNIPROT:P07631; UNIPARC:UPI000017581B
 C:Superfamily: alcohol sulfotransferase
 K:Keywords: liver

Alignment Scores:

Alignment Scores:		
Pred. NO.:	1.62e-17	282
Score:	338.00	89
Percent Similarity:	48.9%	47
Best Local Similarity:	32.8%	Conservative: 89
Query Match:	7.7%	Mismatches: 112
Indels:		30
Gaps:		6
DP:	2	

US-10-768-158-1 (1-2419) x A26136 (1-282)

QY	69	AGCAAGTACTTCGATTCCATGGCGTGGCGGTGCGG-----CCCTTCGCGCGGGGAAG	122
Db	3	SerAspTyrAsenTrpPheGluGlyIleProPheProAlaIleSerTyrGlnArgGluLeu	22
QY	123	ATGAGGAGATCGCCAC--TTCCCGGTGGCGCCACGACGAGTGTGGATCGTCACTAC	179
Db	23	LeuGluAspIleArgAsnLysPheValValLysGluAspLeuLeuLeuThrTyr	42
QY	180	CCCAAGTCCGACACAGCTCTCTCGAGAGGTGGTCTACTTGGTGAGCCAGCGGCGCTGAC	239
Db	43	ProLysSerGlyThrAsenTrpLeuAsnGluLeuValCysLeuIleGlnThrLysGlyAsp	62
QY	240	CCCGATGAGATC-----GGCTTGATGAACATCGACGAGCAGCTCCCGTC	284
Db	63	ProLysTrpIleGlnSerCysProPheGlyThrValTyrProAspGluIleGluTrpIle	82
QY	285	CTGAGTACCCACAGCCCGGCTGGACATCATCAAGAACTGACCTCTCCCGCCTCATC	344
Db	83	PheArgAsnAsnHisGlyGly-----ProArgLeuIle	93
QY	345	AAGACCCACCTCGCCCTACCGCTTCTGCCTCTGACCTCCAAATGGAGACTCCAAAGTC	404
Db	94	ThrSerHisLeuProIleHisLeuPheSerLysSerPhePheSerSerLysAlaLysAla	113
QY	405	ATCTATATGGTTCGAACCCAGATCTGGTGGTCTTATTATCAGTTCCACCGCTCT	464
Db	114	IleTyrLeuMetArgAsnProArgAspIleLeuValSerGlyTyrPhePheTrpGlyAsn	133
QY	465	CTGCGGACCATGAGCTACCGAGGCACCTTCAAGAATTCTGCGGAGGTTTATGAATGAT	524
Db	134	ThrAsnLeuValLysAsnProGlySerLeuGlyThrTyrPheGluTrpPheLeuGlnGly	153
QY	525	AAGCTGGGCTACCGGCTCTCGTTTCAGCACGTCGACGAGTCTGGAGCACCACATGGAC	584
Db	154	AsnValLeuPheGlySerTrpPheGluHisValArgGlyTrpLeuSerMetArgGluTrp	173
QY	585	TCGAACGTGCTTTTCTCAAGTATGAACATGCACTGGGACCTGGTGACGATGTTGGGAG	644
Db	174	AspAsnPheLeuValLeuTyrTrpGluAspMetLysLysAspThrMetGlyThrIleLys	193
QY	645	CAGCTGGCCAGATCTCTGGGGTGTCCTGTGACAAAGGCCACAGTGAAGCCCTGACGGAG	704
Db	194	LysIleCysAspPheLeuGlyLysAsnLeuGlyProAspGluLeuAspLeuLeuLys	213
QY	705	CACCTGC-----CACCAGCTGGTGGACACG	728

Db 214 TyrSerSerPheGlnAlaMetLysGluAsnAsnMetSerAsnTyrSerLeuLeuLysGlu 233

QY 729 TGCTGCAACGCTGAGGCCCTGCCCGTGGCCGG---GGAAGAGTTGGCTGTGGAAAGGAC 785

Db 234 AspProIleLeuThrGlyLeuLysLeuMetArgLysGlyThrThrGlyAspTrpLysAsn 253

QY 786 ATCTTCACCGTCTCCATGAATGAGAAGTTTGACTTGCTATATAACAGAAGATG 839

Db 254 HisPheThrValAlaGlnAlaGluAlaPheAspLysValPheGlnGluLysMet 271

RESULT 40

B40216

Flavonol 3'-sulfotransferase - Flaveria chloraefolia

C:Species: Flaveria chloraefolia

C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004

C:Accession: B40216

R:Varin, L.; Deluca, V.; Ibrahim, R.K.; Brieson, N.

Proc. Natl. Acad. Sci. U.S.A. 89, 1286-1290, 1992

A:Title: Molecular characterization of two plant flavonol sulfotransferases.

A:Reference number: A40216; MUID:92159034; PMID:1741382

A:Accession: B40216

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: nucleic acid

A:Residues: 1-311 <VAR>

A:Cross-references: UNIPROT:P52836; UNIPARC:UPI000012A3FC; GB:M84135; NID:g168166; PID:g

A:Experimental source: terminal bud

A:Note: sequence extracted from NCBI backbone (NCBIP:82218)

C:Superfamily: alcohol sulfotransferase

Alignment Scores:

Pred. No.: 7,53e-15 Length: 311

Score: 303.00 Matches: 82

Percent Similarity: 47.9% Conservative: 57

Best Local Similarity: 28.3% Mismatches: 95

Query Match: 6.9% Indels: 56

DB: 2 Gaps: 11

US-10-768-158-1 (1-2419) x B40216 (1-311)

QY 108 TTCTGCGCGGGAAGATGAGGAGATCGCCAACTTCCCGTGGCGCCAGCGAGTGTGG 167

Db 34 PheLeuGluGlyArgIleLeuSerGluGlnLysPheLysAlaHisProAsnAspValPhe 53

QY 168 ATCGTCACCTACCCCAAGTCGCGCACAGCTTGCTGCAGGAGGTGTCTACTTGTGTGAGC 227

Db 54 LeuAlaSerTyrProLysSerGlyThrThrTrpLeuLysAlaTrpIleCysIleIleThr 73

QY 228 CAGGGCGCTGACCCCGATGATCGGC---TTGATGAATC-----GACGACAGCTC 278

Db 74 ArgGluLysPheAspAspSerThrSerProLeuLeuThrThrMetProHisAspCysIle 93

QY 279 CCGGTCTCTGGAGTACCCACAGCGCGGCTGGACATCATCAAGGAA----- 323

Db 94 ProLeuLeuGlu-----LysAspLeuGluLysIleGlnGluAsnGlnArgAsnSer 110

QY 324 CTGACCTCTCCCGCTCATCAAGACCCACTGCTCCCTTCTGCTTCTGCTGTGACCTC 383

Db 111 LeuTyrThrPro-----IleSerThrHisPheHisTyrLysSerLeuProGluSerAla 128

QY 384 CACAATGGAGACTCCCAAGTCTATATATGGCTCGCAACCCAGGATCTGTGTGTCT 443

Db 129 ArgThrSerAsnCysLysIleValTyrIleTyrArgAsnMetLysAspValIleValSer 148

QY 444 TATTATCAGTTCCACCGCTCTCTGCGGACCATGAGC-----TACCGAGGC 488

Db 149 TyrTyrHisPheLeuArgGlnIleValLysLeuSerValGluAlaProPheGluGlu 168

QY 489 ACCTTTCAAGAAATTCGCCGAGGTTTATCAATGATAGCTGGGCTACGGCTCTCGTTT 548

Db 169 AlaPheAspGluPheCysGlnGlyIleSerSer-----CysGlyProTyrTrp 184

QY 549 GAGCAGTCGAGGAGTTCTGGGAGCACCGCATGGACTCGAAC-----GTGCTTTTCTC 602

Db 185 GluHisIleLysGlyTyrTrpLysAlaSerLeuGluLysProGluLeuPheLeuPheLeu 204

QY 603 AAGTATGAGACATCATCGGACCTGGTGTGACATGGTGGAGCAGCTGGCCAGATTCTCTG 662

Db 205 LysTyrGluAspMetLysLysAspProValProSerValLysLysLeuAlaAspPheIle 224

QY 663 GGG-----GTGTCC 671

Db 225 GlyHisProPheThrProLysGluGluGluAlaGlyValIleGluAspIleValLysLeu 244

QY 672 TGTGACAAGCCGACGCTGGAAGCCCTG-----ACGGAG 704

Db 245 CysSerPheGluLysLeuSerSerLeuGluValAsnLysSerGlyMethHisArgProGlu 264

QY 705 CACTGCCACACGCTGGTGGACGACATCTTCCACGCTCTCCATGAATGAGAGTTTACTTGGTG 824

Db 265 GluAlaHisSerIleGluAsnArg-----LeuTyrPheArgLysGly 278

QY 765 AGAGTTGGCTGTGGAGGACATCTTCCACGCTCTCCATGAATGAGAGTTTACTTGGTG 824

Db 279 LysAspGlyAspTrpLysAsnTyrPheThrAspGluMetThrGlnLysIleAspLysLeu 298

QY 825 TATAAACAGAGATGGAAAGTGTGACCTC 854

Db 299 IleAspGluLysLeuGlyAlaThrGlyLeu 308

RESULT 41

T47448

sulfotransferase-like protein - Arabidopsis thaliana

N:Alternate names: protein T14D3.20

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004

C:Accession: T47448

R:Jordan, N.; Bangert, S.; Wiedelmann, R.; Voss, H.; Unseld, M.; Mewes, H.W.; Lemcke, K.

submitted to the Protein Sequence Database, February 2000

A:Reference number: Z24467

A:Accession: T47448

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-329 <JON>

A:Cross-references: UNIPROT:O9M1V1; UNIPARC:UPI000009DA7F; EMBL:AL138649

A:Experimental source: cultivar Columbia; BAC clone T14D3

C:Genetics:

A:Map position: 3

A:Note: T14D3.20

C:Superfamily: alcohol sulfotransferase

Alignment Scores:

Pred. No.: 9,08e-15 Length: 329

Score: 302.00 Matches: 89

Percent Similarity: 45.2% Conservative: 61

Best Local Similarity: 26.8% Mismatches: 108

Query Match: 6.9% Indels: 74

DB: 2 Gaps: 15

US-10-768-158-1 (1-2419) x T47448 (1-329)

QY 27 GAGAGCGAGCGCGAGACCCCGGGAGTTCGAGAGCAAGTACTTCGAGTTC 86

Db 18 GluSerLysThrLeuIleSerSerLeuProSerAspLysAsnSerThr----- 33

QY 87 CATGCGCTCGCGCTG-----CCGCCCTTCTGCGCGGG 119

Db 34 ---GlyValAsnValCysLysTyrGlnGlyCysTyrTrpTyrThrProIleLeuGlnGly 52

QY 120 AGATGGAGGAGATCCCAACTTCCGGTGGCGGCCGAGCGAGTGTGGTCTGCACCTAC 179

Db 53 ValLeuAsnPheGlnLysAsnPheLysProGlnAspThrAspIleValAlaSerPhe 72

QY 180 CCCAAGTCGCGCACGAGCTTGTCTGAGGAGGTGGTCTAC---TTGCTGAGCCAGGCGGCT 236

Db 73 ProLysCysGlyThrThrTrpLeuLysAlaLeuThrPheAlaLeuValArgSerLys 92

```
QY 237 GACCCCGATGAGATCGGCTTGATGAACATCGACGAGCAGCTCCCGTCTCTG---CAGTAC 293
Db 93 HisProSerHis-----AspAspHisHisProLeuLeuSerAspAsn 106
QY 294 CCA-----CAGCCGGCCTG-----GAC 311
Db 107 ProHisValLeuSerProSerLeuGluMetTyrLeuTyrLeuCysSerGluAsnProAsp 126
QY 312 ATCATCAAGGAACACTGACCTCTCCCGCTCATCAAGAGCCACCTGCCCTACCGCTTTCTG 371
Db 127 LeuThrLysPheSerSerSerArgLeuPheSerThrHisMetProSerHisThrLeu 146
QY 372 CCTCTGACCTCCAAATGGAGACTCCAAAGTCATCTATATGCTCGCAACCCCAAGGAT 431
Db 147 GlnGluGlyLeuLysGlySerThrCysLysIleValTyrMetSerArgAsnValLysAsp 166
QY 432 CTGGTGGTCTCTATTATCATGTTCCACCGCTCTCTCGGACG-----ATG 476
Db 167 ThrLeuValSerTyrTrpHisPheCysLysLysGlnThrAspAsnIleIleSer 186
QY 477 AGCTACCGAGGACCTTTCAAGATTCTGCCG-----AGTTTATGAATGATAAGCTG 530
Db 187 SerValGluAspThrPheGluMetPheCysArgGlyValAsnPhe-----201
QY 531 GGCTACGGCTCTGTTTGTGACGACGTGACGAGTCTGGGACACCGCATGGAC-----584
Db 202 ---PheGlyProPheTrpAspHisValLeuSerTyrTrpArgLysSerLeuGluAspPro 220
QY 585 TCGAAGCTGCTTTTCTCAAGTATGAAGACATCATCGGAGCTTCTGGGACACCGCATGGAG 644
Db 221 AsnHisValLeuPheMetLysPheGluGluMetLysGluGluProArgGluGlnIleLys 240
QY 645 CAGCTGGCCAGATTCTCGGGGTGCTCTGTGACAGGCCACCGCTGGAAGCC---CTGACG 701
Db 241 ArgLeuAlaGluPheLeuGlyCysLeuPheThrLysGluGluGluSerGlyLeuVal 260
QY 702 GAGCACTGCCACGAGCTGTGGACCACTGCTGC-----AACGCTGAGCCCTGCCCGTG 755
Db 261 Asp-----GluIleIleAspLeuCysSerLeuArgAsnLeuSerLeuGluIle 277
QY 756 GGC-----CGGGA 764
Db 278 AsnLysThrGlyLysLeuHisSerThrGlyArgGluAsnLysThrPheArgLysGly 297
QY 765 AGAGTTGGGCTGGGAAGACATCTTCACCGTCTCCATGAATGAGAAGTTGACTGGTG 824
Db 298 GluValGlyAspTrpLysAsnTyrLeuThrProGluMetGluAsnLysIleAspMetIle 317
QY 825 TATAACAGAAAGATGGGAAGTGTGACCTCACGTTT 860
Db 318 IleGlnGluLysLeuGlnAsnSerGlyLeuLysPhe 329

RESULT 42
B84452
probable steroid sulfotransferase [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: B84452
R.;Lin. X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
euss, D.; Nierman, W.C.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
M.; Koo, H.; Moffat, K.S.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-324 <STO>
A;Cross-references: UNIPROT:Q9ZPQ5; UNIPARC:UPI00000A21B0; GB:A8002093; NID:g4406769; P
C;Genetics:
A;Gene: At2g03770
A;Map position: 2
C;Superfamily: alcohol sulfotransferase
```

```
Alignment Scores: 1.99e-14 Length: 324
Pred. No.: 297.50 Matches: 84
Score: 43.3% Conservative: 57
Percent Similarity: 25.8% Mismatches: 124
Best Local Similarity: 6.8% Indels: 61
Query Match: 2 Gaps: 10
DB:

US-10-768-158-1 (1-2419) x B84452 (1-324)
QY 21 ATGGCGGAGAGCGAG-----GCCGAGACCCCCAGCACC 53
Db 12 LeuValGluSerGluLeuValGlnCysGluGluLeuLeuSerSerLeuProArgAsp 31
QY 54 CCGGGGAGTTTCGAGAGCAAGTACTTCAGTTCCATGCGTGGCGCTGCCGCC---TTC 110
Db 32 ArgSerValPheAlaGluTyrLeuTyrGlnTyrGlnGlyPheTrpTyrProProAsnLeu 51
QY 111 TGCCCGCGGAGATCGAGAGATCCCAACTTCCCGTCCGCCGAGCGAGCTGTGGATC 170
Db 52 LeuGluGlyValLeuTyrSerGlnLysHisPheGlnAlaArgAspSerAspIleValLeu 71
QY 171 GTCACCTACCCCAAGTCCCGCAGCAGCTTGTCTGCGAGAGTGGTCTACTTGTGTGAGCCAG 230
Db 72 AlaSerIleProLysSerGlyThrThrTrpLeuLysSerLeuValPhe-----87
QY 231 GCGCTGACCCCGATGAGATCGGCTTGTATGAACATCGACGAG---CAGCTCCCGTCTCTG 287
Db 88 -----AlaLeuIleHisArgGlnGluPheGlnThrProLeuVal 100
QY 288 GAGTACCACACAG-----CCGGCGCTGGAC 311
Db 101 SerHisProLeuLeuAspAsnAsnProHisThrLeuValThrPheIleGluGlyPheHis 120
QY 312 ATCATCAAGAACTACCTCTCCCGCTCATCAAGAGCCACCTGCCCTTACCGCTTCTG 371
Db 121 LeuHisThrGlnAspThrSerProArgIlePheSerThrHisIleProValGlySerLeu 140
QY 372 CCTCTGACCTCCAAATGGAGACTCCCAAGTCTATATGCTGCGCAACCCCAAGGAT 431
Db 141 ProGluSerValLysAspSerSerCysLysValValTyrCysArgAsnProLysAsp 160
QY 432 CTGGTGGTCTTATATATACAGTTCCACCGCTCTCTGCGGACCATGAGCTACCGAGC---488
Db 161 AlaPheValSerLeuThrHisPheMetLysAsnLeuIleValLysGluMetValGlyCys 180
QY 489 ACCTTTCAAGAAATTCGCGGAGGTTTATGAATGATGAAGCTGGCTACCGCTCTGGTTT 548
Db 181 ThrMetGluGluMetValArgPhePheCysArgGlySerSerIleTyrGlyProPheTrp 200
QY 549 GAGCAGCTGACGAGTTCGGAGACACCGCATGGACTCG-----AACGTGCTTTTCTC 602
Db 201 AspHisValLeuGlnTyrTrpLysGluSerArgGluAsnProLysLysValMetPheVal 220
QY 603 AGATATGAAGACATCATCGGACCTGGTGACGATGGTGGAGCAGCTGGCCAGATTCCTG 662
Db 221 MetTyrGluGluMetArgGluGlnProGlnGluTrpValMetArgIleAlaGluPheLeu 240
QY 663 GGGGTCTCTGTGCAAGGCCCGCTGGAA---GCCCTGACGAGGACACTGCCACCAAGCTG 719
Db 241 GlyTyrSerPheThrGluGluGluIleGluAenglyValLeuGluAspIleIleLysLeu 260
QY 720 GTGACCACTGTGTGAACCGCTGAGGCCCTGCC---752
Db 261 -----CysSerLeuGluAsnLeuSerLysLeuValAsnGluLysGlyLys 276
QY 753 -----GTGGCGCGGGAAGAGTTGGCTGTGGAAG 782
Db 277 LeuLeuAsnGlyMetGluThrLysAlaPhePheArgLysGlyGluIleGlyLysTrpArg 296
QY 783 GACATCTTCACTCTCCATGAATGAGAAGTTTGACTTGTGTATAAACAGAGATGGGA 842
Db 783 GACATCTTCACTCTCCATGAATGAGAAGTTTGACTTGTGTATAAACAGAGATGGGA 842
```

Db 297 AspThrLeuThrProLeuLeuAlaGluGluIleAspLysThrThrThrLysGluLysLeuIle 316
 QY 843 AAGTGTGACCTCACGTTT 860
 Db 317 GlySerAspPheArgPhe 322
 RESULT 43
 T47447
 sulfotransferase-like protein - Arabidopsis thaliana
 N:Alternate names: protein T14D3.10
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
 C:Accession: T47447
 R:Jordan, N.; Bangert, S.; Wiedelmann, R.; Voss, H.; Unseld, M.; Mewes, H.W.; Lemcke, K.
 submitted to the Protein Sequence Database, February 2000
 A:Reference number: Z24467
 A:Accession: T47447
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-323 <J0R>
 A:Cross-references: UNIPROT:Q9M1V2; UNIPARC:UPI00000A990C; EMBL:AL138649
 A:Experimental source: cultivar Columbia; BAC clone T14D3
 C:Genetics:
 A:Map position: 3
 A:Note: T14D3.10
 C:Superfamily: alcohol sulfotransferase

Alignment Scores:
 Pred. No.: 1-93e-13 Length: 323
 Score: 284.50 Matches: 74
 Percent Similarity: 45.7% Conservative: 58
 Best Local Similarity: 25.6% Mismatches: 90
 Query Match: 6.5% Indels: 67
 DB: 2 Gaps: 11

US-10-768-158-1 (1-2419) x T47447 (1-323)

QY 138 AACTTCCGGTGGGCCCCAGCGAGTGTGTGATGTGCACCTACCCCAAGTCCGCGACACG 197
 Db 54 SerPheLysProGlnAspThrAspIleValAlaSerPheProLysCysGlyThrThr 73
 QY 198 TTGTCGACGAGTGTCTAC---TTGGTGAGCCAGGCGCTGACCCCGATGATGATCGGC 254
 Db 74 TrpLeuLysAlaLeuThrPheAlaLeuLeuHisArgSerLysGlnProSerHis----- 91
 QY 255 TTGATGAACATCGACGAGCAGCTCCCGTCCCTG----- 287
 Db 92 -----AspAspHisProLeuLeuSerAsnAsnProHisValLeuValPro 107
 QY 288 -----GAGTACCCACAGCCGCGGCTGGACATCATC 317
 Db 108 TyrPheGluIleAspLeuTyLeuArgSerGluAsnPro-----AspLeuThr 123
 QY 318 AAGGAACACTGCTCCCGCTCATCAGAGCCACTGCTGCTTCTGCGCTCT 377
 Db 124 LysPheSerSerProArgLeuPheSerThrHisValProSerHisThrLeuGlnGlu 143
 QY 378 GACCTCCCAATGGAGACTCCAGGTCATCATATATGCTCGCAACCCAGGATCTGGG 437
 Db 144 GlyLeuLysGlySerThrCysLysIleValTyLeSerArgAsnValLysAspThrLeu 163
 QY 438 GTGCTCTATTATCAGTTCACCGCTCTCTCGGAGC-----ATGAGCTAC 482
 Db 164 ValSerTyTrpHisPheThrLysLysGlnThrAspGluLysIleIleSerSerPhe 183
 QY 483 CGAGGCACCTTTCAGAAATTCGCGGAGGTTTATGAATGATAGTCTGGGC----- 533
 Db 184 GluAspThrPheGluMetPheCysArg-----GlyValSerIle 196
 QY 534 TAGGGCTCTGGTTTTCAGACGTCGAGGAGTCTGGGACACCGCATGAC-----TCG 587
 Db 197 PheGlyProPheTrpAspHisValLeuSerTyTrpArgGlySerLeuGluAspProAsn 216

QY 588 AACGTGCTTTTCTCAAGTATGAAGCATGTCCGGACCTCGTGACGATGTGTGAGCAG 647
 Db 217 HisValLeuPheMetLysPheGluGluMetLysAlaGluProArgAspGlnIleLysLys 236
 QY 648 CTGGCCAGATTCTCGGGGTGCTGTGACAAGGCCAGCTGGAAGCCCTGACGGAGCAC 707
 Db 237 PheAlaGluPheLeuGly-----CysProPheThrLysGluGluGluSerGlySer 254
 QY 708 TCCCAACAGCTGGTGACAGCTGCTGC-----AACGTGAGGCCCTGCGCTGGGC--- 758
 Db 255 ValAspGluIleIleAspLeuCysSerLeuArgAsnLeuSerSerLeuGluIleAsnLys 274
 QY 759 -----CGGGGAAGATTGGG 773
 Db 275 ThrGlyLysLeuAsnSerGlyArgGluAsnLysMetPhePheArgLysGlyGluValGly 294
 QY 774 CTGTGAAGGACATCTTACCGTCTCCATGAATGAGAAAGTTTTCAGCTTGGTGTATAAACAG 833
 Db 295 AspTrpLysAsnTyLeuThrProGluMetGluAsnLysIleAspMetIleIleGlnGlu 314
 QY 834 AAGATGGGAAGTGTGACCTCACGTTT 860
 Db 315 LysLeuGlnAsnSerGlyLeuLysPhe 323

RESULT 44
 T07833
 probable steroid sulfotransferase (BC 2.8.2.15) 3 - rape
 C:Species: Brassica napus (rape)
 C:Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 09-Jul-2004
 C:Accession: T07833
 R:Richard, M.; Nicolle, L.; Varin, L.
 submitted to the EMBL Data Library, April 1997
 A:Reference number: Z16161
 A:Accession: T07833
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-325 <R1C>
 A:Cross-references: UNIPROT:O82410; UNIPARC:UPI00000AA137; EMBL:AF000307; NID:G3420007;
 C:Genetics:
 A:Gene: ST3
 C:Superfamily: alcohol sulfotransferase
 C:Keywords: sulfotransferase

Alignment Scores:
 Pred. No.: 4.25e-13 Length: 325
 Score: 280.00 Matches: 82
 Percent Similarity: 46.1% Conservative: 60
 Best Local Similarity: 26.6% Mismatches: 118
 Query Match: 6.4% Indels: 48
 DB: 2 Gaps: 12

US-10-768-158-1 (1-2419) x T07833 (1-325)

QY 45 CCAGACCCCGGGGAGTTCCAGAGCAAGTACTTCGAGTTTCATGGC---GTGCGGCTG 101
 Db 28 ProSerGluLysGlyTrpLeuValSerGlnIleTyTrpGlnPheGlnGlyArgTrpHisThr 47
 QY 102 CGCGCTTCTGCGCGGGAAGATGAGGAGATCGCAACTCCCGTGGCGGCCACCGAC 161
 Db 48 GluAlaLeuLeuGlnGlyIleLeuThrCysGlnLysHisPheLysAlaLysAspSerAsp 67
 QY 162 GTGTGGATGCTCACCTACCCCAAGTCCGCGCACAGCTTGTCTGCAGAGAGTGTCTACTTG 221
 Db 68 IleIleLeuValThrAsnProLysSerGlyThrThrTrpLeuLysSerLeuValPheAla 87
 QY 222 -----GTGAGCCAGGGCGCTGACCCCGATGATCGGCTG 257
 Db 88 LeuIleAsnArgHisLysPheProValSerSerGlyAspHisPro-----LeuLeuVal 105
 QY 258 ATGAACATCGACGACGCTCCCGTCTCTGGAG-----TACCACAGCCGGCGCTG 308
 Db 106 ThrAsnProHisLeuLeuValProPheMetGluGlyValTyTrpGluSerProAspPhe 125


```
QY 309 GACATCATCAAGAACTGACCTCTCCCGCCTCATCAAGAGCCACTCGCCCTACCGCTTT 368
Db 126 Asp-----PheSerLeuLeuProPheProArgLeuMetAsnThrHisIleSerHisLeuSer 144
QY 369 CTGCGCTCTGACCTCCACAATGGAGACTCCAAAGTCACTATATATGGCTGCGCAACCCCAAG 428
Db 145 LeuProGluSerValLysSerSerCysGlnIleValTyrcysArgAsnProLys 164
QY 429 GATCTGGTGGTCTCTATTATCAGTTTCACCCCTCTCTG-----CGGACCATGAGC 479
Db 165 AspMetPheValSerLeuTrpHisPheGlyLysPheGlyLysLeuAlaProGlnGluThrAlaAsp 184
QY 480 TACCAGGACCTTTCAGAAATCTGCGGAGGTTTATGAATGATAAGCTGGGCTACGCG 539
Db 185 Tyr-----ProLeuGluLysAlaValGluAlaPheCysGlnGlyPheIlealagly 202
QY 540 TCTCTGTTTGGACAGCTGAGGAGTTCTGGGAGCACCGCATCGAC-----TCGAACGCTG 593
Db 203 ProPheTrpAspHisValLeuGluTyrrTrpTyrrAlaSerLeuGluAsnProAsnLysVal 222
QY 594 CTTTTCCTCAAGTATGAACATGATCGGAGCTCGGACCTGGTGACATGGTGAGAGCTGGCC 653
Db 223 LeuPheValThrTyrrGluGluLeuLysLysGlnThrGluValGluValLysArgIleAla 242
QY 654 AGATTCTGGGGTCTCTGTGACAAAGGCCAGCTGGAGCCCTGACGAGCACTGCCAC 713
Db 243 GluPheIleGly-----CysGlyPheThrAlaGluGluGluValSerGlu----- 257
QY 714 CAGCTGGTGGAGCAGCTGCTGCAACGCTGAGGCC-----CGGGAGAGTTGGGCTG 746
Db 258 -----IleValLysLeuCysSerPheGluSerLeuSerArgLeuGluValAsnArgGln 275
QY 747 -----CTCCCGCTGGCC-----CGGGAGAGTTGGGCTG 776
Db 276 GlyLysLeuProAsnGlyIleGluThrAsnAlaPheArgLysGlyGluIleGly 295
QY 777 TGAAGGACATCTCACCTCTCCATGATGAGAGTTTGACTGGTGTATAAACAGAG 836
Db 296 TrpArgAspThrLeuSerGluSerLeuAlaAspAlaIleAspArgThrThrGluGluLys 315
QY 837 ATGGGAAAGTGTGACCTCAGCTTT 860
Db 316 PheGlyGlySerGlyLeuLysPhe 323
RESULT 45
A:4452
probable steroid sulfotransferase [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: A84452
R.Jolin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.B.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: A84452
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-326 <STO>
A:Cross-references: UNIPROT:P52839; UNIPARC:UPI00000484B4; GB:AE002093; NID:G4406767; PI
A:Gene: At2g03760
A:Map position: 2
C:Superfamily: alcohol sulfotransferase
```

```
Alignment Scores:
Pred. No.: 5.06e-13 Length: 326
Score: 279.00 Matches: 87
Percent Similarity: 44.6% Conservative: 58
Best Local Similarity: 26.8% Mismatches: 128
Query Match: 6.3% Indels: 52
DB: 2 Gaps: 13
```

```
US-10-768-158-1 (1-2419) x A84452 (1-326)
QY 6 GCGGAGCGCGCGGATGCGGAGAGCGAGCGCCGAGACCCCGAGCACC----- 56
Db 12 GlyAspGluAspLeuThrGlnGluThrArgAlaLeuIleSerSerLeuProLysGluLys 31
QY 57 GGGAGTTTCGAGAGCAAGTACTTCAGTTTCATCGCGTG---CGGCTGCCCGCTTCGTC 113
Db 32 GlyTrpLeuValSerGluIleTyrrGluPheGlnGlyLeuTrpHisThrGlnAlaIleLeu 51
QY 114 CCGCGGAAGATGCGAGGAGATCGCAACTTCGCGCGCGCGCCAGCGAGCTGTGATCGTC 173
Db 52 GlnGlyIleLeuIleCysGlnLysArgPheGluAlaLysAspSerAspIleIleLeuVal 71
QY 174 ACCTACCCCAAGTCCGCGCACCGAGTTGCTGAGAGGTGGTCTACTCTGGT----- 224
Db 72 ThrAsnProLysSerGlyThrThrTrpLeuLysAlaLeuValPheAlaLeuLeuAsnArg 91
QY 225 -----AGCCAGGCGCTCACCCGATGAGATCGGCTTGTGAACATC 266
Db 92 HisLysPheProValSerSerGlyAsnHisPro-----LeuLeuValThrAsnPro 109
QY 267 GACGAGCAGCTCCCGCTCTCGAG-----TACCCACAGCGCGGCTCGACATCATC 317
Db 110 HisLeuLeuValProPheLeuGluGlyValTyrrTyrrGluSerProAspPheAsp---Phe 128
QY 318 AAGGAATGACCTCTCCCGCTCATCAAGAGCACCTGCGCTTCTTCTGCGCTCT 377
Db 129 SerSerLeuProSerProArgLeuMetAsnThrHisIleSerHisLeuSerLeuProGlu 148
QY 378 GACCTCCCAATGAGAGCTCCAAGTCACTATATGCTCGCAACCCCAAGAGTGGTG 437
Db 149 SerValLysSerSerCysLysIleValTyrrCysCysArgAsnProLysAspMetPhe 168
QY 438 GTGCTTATTATCAGTTCACCGCTCTCTG-----CGACCATGAGTACCGAGGC 488
Db 169 ValSerLeuTrpHisPheGlyLysLeuAlaProGluGluThrAlaAspTyr----- 186
QY 489 ACCTTTCAAGAAATTCGCGGAGTTTATGAATGATAGCTGGGCTACGGCTCTGTTT 548
Db 187 ProIleGluLysAlaValGluAlaPheCysGluGlyLysPheIleGlyGlyProPheTrp 206
QY 549 GAGCAGCTGCGAGGATCTCG-----CAGCACCGCATGGACTCGAACGCTGTTTCTC 602
Db 207 AspHisIleLeuGluTyrrTrpTyrrAlaSerArgGluAsnProAsnLysValLeuPheVal 226
QY 603 AAGTATGAAGACATCATCGCGGACCTGGTGAGAGCTGGGAGCTGGCCAGATTCCTG 662
Db 227 ThrTyrrGluGluLeuLysGlnThrGluValGluMetLysArgIleAlaGluPheLeu 246
QY 663 GGGGTGCTCTGTGCAAGCGCCAGCTGGAAGCCCTGAGCGAGCACTGCCACCGCTGGTG 722
Db 247 -----GluCysGlyPheIleGluGluGluValArgGlu-----Ile 259
QY 723 GACCACTGCTCGAAGCTGAGGCC-----CTG 749
Db 260 ValLysLeuCysSerPheGluSerLeuValAsnLysGluGlyLysLeu 279
QY 750 CCGCTGGGC-----CGGGAAGAGTTGGGCTGGGAAGAC 785
Db 280 ProAsnGlyIleGluThrLysThrPhePheArgLysGlyGluIleGlyTyrrArgAsp 299
QY 786 ATCTTTCACCGCTCCATGAATGAGAAGTTTGACTTTGGTGTATAAACAGAAGATGGAAG 845
Db 300 ThrLeuSerGluSerLeuAlaGluIleAspArgThrIleGluGluLysPheLysGly 319
QY 846 TGTGACCTCAGCTTT 860
Db 320 SerGlyLeuLysPhe 324
RESULT 46
E86319
```


probable flavonol sulfotransferase [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
 C:Accession: E86319
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.; Jensen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; C.A.; Li, J.H.; Lin, Y.; Lin, X.; Liu, Z.A.; Luros, J.S.; Maity, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: E86319
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-346 <STO>
 A:Cross-references: UNIPROT:Q9FZ80; UNIPARC:UPI000000C631; GB:AE005172; NID:g9795597; PI
 A:Map position: 1
 A:Superfamily: alcohol sulfotransferase

Alignment Scores:
 Pred. No.: 5.6e-13 Length: 346
 Score: 278.50 Matches: 83
 Percent Similarity: 43.5% Conservative: 62
 Best Local Similarity: 24.9% Mismatches: 119
 Query Match: 6.3% Indels: 69
 DB: 2 Gaps: 12

US-10-768-158-1 (1-2419) x E86319 (1-346)

QY 21 ATGGCGGAGAGAGAGCCGAG---ACCCCGACAGCCCGGGGAGTTTCGAG-----AGC 71
 Db 11 ValSerGluSerAsnHisGluLeuAlaSerSerProSerGluPheGluLysAsnGln 30
 QY 72 AAGTACTTCGAGTTCATCGCGTGGCGTGGCG-----104
 Db 31 LysHisTyrGlnGluIleAlaThrLeuProHisLysAspGlyTrpArgProLysAsp 50
 QY 105 -----CCCTTCGCGGGGAGAGATGAG 128
 Db 51 ProPheValGluTyrGlyGlyHisTrpTrpLeuGlnProLeuLeuGluGlyLeuLeuHis 70
 QY 129 GAGATCGCCAACTTCCCGTGGCGCCGACGAGTGTGATCGTCACCTACCCCAAGTCC 188
 Db 71 AlaGlnLysPhePheLysAlaArgProAsnAspPheValCysSerTrpProLysThr 90
 QY 189 GGCACGAGCTTCTGCAGGAGTGTCTACTTGTGAGCGAGCGGCTGACCCCGAT--- 245
 Db 91 GlyThrThrTrpLeuLysAlaLeuThrPheAlaIleAlaAsnArgSerLysPheAspVal 110
 QY 246 -----GAGATCGGTTGATGAACATCGACGAGCAGCTCCCGGTCTCGAG----- 290
 Db 111 SerThrAsnProLeuLysArgAsnProHisGluPheValProTyrIleGluIleAsp 130
 QY 291 TACCCA---CAGCGGCGCTGCACATCATCAGGAACTGACCTCTCCCGCTCATCAAG 347
 Db 131 PheProPhePheProSerValAspValLeuLysAsp---GluGlyAsnThrLeuPheSer 149
 QY 348 AGCCACCTGCGCTACCGCTTTCTGCGCTTCTGACCTCCACAAATCGAGACTCCAAAGGTCTATC 407
 Db 150 ThrHisIleProTyrAspLeuLeuProGluSerValValLysSerGlyCysLysIleVal 169
 QY 408 TATATGGTTCGCAACCCCAAGGATCTGGTGTGTATTATATGATTCACCGCTCTCTG 467
 Db 170 TyrIleTrpArgAspProLysAspThrPheValSerMetTrpThrPheAlaHisLysGlu 189
 QY 468 CGGACC-----ATGAGCTACCCGAGGACCTTTCAGAAATCTTCGCGGAGG 512
 Db 190 ArgSerGlnGlnGlyProValValSerIleGluGluAlaPheAspLysTyrCysGlnGly 209

QY 513 TTTATGAATGATAAGCTGGCTACGGCTCCCTGGTTTGAGCAGCTGCAGGAGTTCGGGAG 572
 Db 210 LeuSer-----AlaTyrGlyProTyrLeuAspHisValLeuGlyTyrTrpLys 225
 QY 573 -----CACCGCATGACTCGAACCTGCTTTTCTCAAGTATGAAGACATGCATCGGAC 626
 Db 226 AlaTyrGlnAlaAsnProAspGlnIleLeuPheLeuLysTyrGluThrMetArgAlaAsp 245
 QY 627 CTGTGTACAGTGTGTGAGCAGCTGCCAGATTCCTGGGGGTGTCTGTGACAAAGCCGAC 686
 Db 246 ProLeuProTyrValLysArgLeuAlaGluPheMetGlyTyrGlyPheThrLysGluGlu 265
 QY 687 CTGGAAGCCCTGACGAGCAGCTGCCACGAGCTGTGGACCACTGTGTGCAACGCTGAGGCC 746
 Db 266 GluGluGly-----AsnValValGluLysValValLysLeuCysSerPheGluThr 282
 QY 747 CTGCCC-----752
 Db 283 LeuLysAsnLeuGluAlaAsnLysGlyLysAspArgGluAspArgProAlaValTyr 302
 QY 753 -----GTGGCGCGGAGAGAGTTCGGCTGTGGAAGACATCTTCAACGCTC 797
 Db 303 AlaAsnSerAlaTyrPheArgLysGlyLysValGlyAspTrpGlnAsnTyrLeuThrPro 322
 QY 798 TCCATGAATGAGAGTTTGTGACTTGTGTATATAACAGAG 836
 Db 323 GluMetValAlaArgIleAspGlyLeuMetGluGluLys 335

RESULT 47

T07831

probable steroid sulfotransferase (EC 2.8.2.15) 1 - rape

C:Species: Brassica napus (rape)

C:Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 09-Jul-2004

C:Accession: T07831

R:Richard, M.; Nicolle, L.; Varin, L.

submitted to the EMBL Data Library, April 1997

A:Reference number: Z16161

A:Accession: T07831

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-323 <RIC>

A:Cross-references: UNIPROT:O82408; UNIPARC:UPI000000A1423; EMBL:AF000305; NID:g3420003;

C:Genetics:

A:Gene: ST1

C:Superfamily: alcohol sulfotransferase

C:Keywords: sulfotransferase

Alignment Scores:
 Pred. No.: 7.82e-13 Length: 323
 Score: 276.50 Matches: 85
 Percent Similarity: 46.5% Conservative: 59
 Best Local Similarity: 27.4% Mismatches: 113
 Query Match: 6.3% Indels: 53
 DB: 2 Gaps: 14

US-10-768-158-1 (1-2419) x T07831 (1-323)

QY 45 CCCAGCACCCCGGGGAGTTCGAGAGCAAGTACTTTCGAGTTCATGGCGTG---CGGCTG 101
 Db 27 ProSerGluLysGlyTrpLeuValSerGlnMetTyrGlnPheGluGlyIleTrpGlnThr 46
 QY 102 CCGCCCTTCTCCCGCGGGAAGATGAGGAGATCGCAACTTCCCGGTGCGGCCCGACGAC 161
 Db 47 GlnAlaLeuValGlnGlyIleValAsnCysGlnLysHisPheGluAlaAsnAspSerAsp 66
 QY 162 GTGTGGATCGTCACTACCCCAAGTCCCGCACCGAGCTTGTGTGAGGAGGTGTCTACTTG 221
 Db 67 ValIleLeuAlaThrLeuAlaLysSerGlyThrThrTrpLeuLysAlaLeuLeuPheAla 86
 QY 222 GTG-----AGCCAGGCGGTGACCCCGATGAGATCGGCTTGATG 260
 Db 87 LeuIleHisArgHisLysPheProValSerGlyLysHisPro-----LeuLeuValThr 104

```

Qy 261 AACATCGACGAGAGCTCCCGGTCCTGAG-----TACCCACAGCCCGGCTGGAC 311
Db 105 AsnProHisSerLeuValProTyrLeuGluGlyAspTyrCysSerSerProGluValAsn 124
Qy 312 ATCATCAGGAACTGACCTCCCGCTCATCAAGAGCCACCTGCTCCGCTTCTG 371
Db 125 ---PheAlaGluLeuProSerProArgLeuMetGlnThrHisLeuThrHisSerLeu 143
Qy 372 CCTCTGACCTCCCAATGGAGACTCCAAAGGTCTATATGGCTCGCAACCCCAAGGAT 431
Db 144 ProValSerIleLysSerSerCysLysIleIleTyrCysArgAsnProLysAsp 163
Qy 432 CTGTGGTGTCTTATATCAGTTCCACCGCTCTCTG-----CGACCATGAGCTAC 482
Db 164 MetPheValSerIleTrpHisPheGlyArgLysLeuAlaProGluLysThrAlaGluTyr 183
Qy 483 -----CGAGGCACCTTCAAGAATTCTGCGG---AGTTTATGAATGAATAGCTGGGC 533
Db 184 ProIleGluThrAlaValAlaAlaPheCysLysGlyLysPheIleGly----- 199
Qy 534 TACGGCTCTGTTTGGAGCAGCTGCGAGAGTTCTGGGAGCACCGCATG-----GACTCG 587
Db 200 ---GlyProPheTrpAspHisValLeuGluTyrTrpTyrGluSerLeuLysAsnProAsn 218
Qy 588 AACGTGCTTTTCTCAAGATGAAGACATGCATCGGACCTGGTGACGATGGTGAGCAG 647
Db 219 LysValLeuPheValThrTyrGluGluLeuLysLysGlnThrGluValGluValysArg 238
Qy 648 CTGGCCAGATTCCTGGGGGTGCTGTCACAAAGGCCAGCTGGAAGCCCTGACGAGCAC 707
Db 239 IleAlaGluPheIleGly-----CysGlyPheThrAlaGluGluValysSerGlu--- 255
Qy 708 TGCCACAGCTGCTGGTGACGAGTGTGCAACGCTGAGGCC----- 746
Db 256 -----IleValLysLeuCysSerPheGluSerLeuSerSerLeuGluValAsn 271
Qy 747 -----CTGCCGCTGGGC-----CGGGAAGAGTT 770
Db 272 ArgGlnGlyLysLeuProAsnGlyIleGluSerAsnAlaPhePheArgLysGlyLeuThr 291
Qy 771 GGGCTGTGAAGGACATCTTACCGCTCTCCATGATGAGAACTTTGACTTGTGTATAAA 830
Db 292 GlyGlyTrpArgAspThrLeuSerGluSerLeuAlaAspValIleAspArgThrThrGlu 311
Qy 831 CAGAAGATGGGAAGTGTGACCTCAGCTT 860
Db 312 GlnLysPheGlyGlySerGlyLeuLysPhe 321

RESULT 48
A40216
Flavonol 4'-sulfoxyltransferase - Flaveria chloraefolia
C:Species: Flaveria chloraefolia
C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 20-Sep-1999
C:Accession: A40216
R:Varin, L.; Deluca, V.; Ibrahim, R.K.; Brissson, N.
Proc. Natl. Acad. Sci. U.S.A. 89, 1286-1290, 1992
A:Title: Molecular characterization of two plant flavonol sulfoxyltransferases.
A:Reference number: A40216; MUID:92159034; PMID:1741382
A:Accession: A40216
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-320 <VAR>
A:Cross-references: UNIPARC:UPI0000175837
A:Experimental source: terminal bud
A:Notes: sequence inconsistent with the nucleotide translation
A:Notes: sequence extracted from NCBI backbone (NCBIN:82216, NCBIIP:82217)
C:Superfamily: alcohol sulfoxyltransferase

Alignment Scores:
Pred. No.: 9.3e-13 Length: 320
Score: 275.50 Matches: 80
Percent Similarity: 49.1% Conservative: 63

```

```

Best Local Similarity: 27.5% Mismatches: 111
Query Match: 6.3% Indels: 37
Db: 2 Gaps: 11

US-10-768-158-1 (1-2419) x A40216 (1-320)

Qy 72 AAGTACTTCGAGTTCCATGGGTGGCGCTTCTGCCGCGGAAGATGAGGAG 131
Db 34 LysTyrGlnAspPheTrpGly-----LeuGlnAsnIleGluGlyAlaIleLeuAla 51
Qy 132 ATCCGCCAACTTCCCGGTGCGGCCCGCAGGACGCTGTGGATCGTCACCTACCCCAAGTCCGCG 191
Db 52 GlnGlnSerPheLysAlaArgProAspAspValPheLeuCysSerTyrProLysSerGly 71
Qy 192 ACCAGCTGTGTCAGGAGGTGCTTACTTGTGTAGCCAGGCGCTGACCCCGATGAGATC 251
Db 72 ThrThrTrpLeuLysAlaLeuAlaTyrAlaIleValThrArgGluLysPheAspGluPhe 91
Qy 252 GGC-----TTGATGAACATCGACGAGCAG---CTCCCGGTCTCTGGAGTACCACAG 299
Db 92 ThrSerProLeuLeuThrAsnIleProHisAsnCysIleProTyrIleGlu----- 108
Qy 300 CCGGCGCTGGACATCATCAAGAACTGACCTCTCCCGC-----CTCATC 344
Db 109 -----LysAspLeuLysLysIleValAspAsnGlnAsnAsnSerCysPheThrProMet 126
Qy 345 AAGAGCCACCTGCGCTTCTGCGCTTCTGCGCTTCTGCGCTTCTGCGCTTCTGCGCTTCTGCGCTTCT 404
Db 127 AlaThrHisMetProTyrHisValLeuProLysSerIleLeuAlaLeuAsnCysLysMet 146
Qy 405 ATCTATATGGTTCGCAACCCCAAGATCTGTGTGTCTTATTATTCAGTTCCTCCACGCTCT 464
Db 147 ValTyrIleTyrArgAsnIleLysAspValIleValSerPheTyrHisPheGlyArgGlu 166
Qy 465 CTGCGGACCATGACGATACCGA---GGCACCTTTCAAGAATTCTCGCGGAGGTTTATGAAT 521
Db 167 IleThrLysLeuProLeuGluAspAlaProPheGluGluAlaPheAspGluPheTyrHis 186
Qy 522 GATAGCTGGGTACGCTCTGTTTTCAGCAGCTGCGAGGATCTGCGGAGCCCGCATG 581
Db 187 GlyIleSerGlnPheGlyProTyrTrpAspHisLeuLeuGlyTyrTrpLysAlaSerLeu 206
Qy 582 GAC-----TCGAAGCTGCTTTTCTCAAGTATGAAGACATGCATCGGAGCTGTGTGACG 635
Db 207 GluArgProGluValIleLeuPheLysLysTyrGluAspValLysLysAspProThrSer 226
Qy 636 ATGTGTGAGCAGCTGCGCAGATTCCTGGGGGTGCTCTGT-----GACAAAG 680
Db 227 AsnValLysArgLeuAlaGluPheIleGlyTyrProPheThrPheGluGluGluLysGlu 246
Qy 681 GCCAGCTGGAAGCCCTGACGAGCAGCTGC-----CACCAGCTGGTG 722
Db 247 GlyValIleGluSerIleLysLeuCysSerPheGluAsnLeuSerAsnLeuGluVal 266
Qy 723 GACCAAGTGTGCAAGCTGAGGCC---CTGCCCGTG-----GGCCGG 761
Db 267 AsnLysSerGlyAsnSerLysGlyPheLeuProIleGluAsnArgLeuTyrPheArgLys 286
Qy 762 GGAAGAGTTGGCTGTGGAAGACATCTTACCGCTCTCATGAATGAGAGTTTACTTG 821
Db 287 AlaLysAspGlyAspTrpLysAsnTyrPheThrAspGluMetThrGluLysIleAspLys 306
Qy 822 GTGTATAACAGACATGGGAAGTGTGACCTC 854
Db 307 LeuIleAspGluLysLeuSerAlaThrGlyLeu 317

RESULT 49
H96768
protein flavonol sulfoxyltransferase E2p9.4 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: H96768
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso

```

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Jensen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
C.A.A.; Li, J.H.; Li, Y.; Lin, X.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; Rizzuto, M.; Rooney, T.; Rowley, D.; Sakano, H.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, K.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: H96768
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-350 <STO>
A:Cross-references: UNIPROT:O9C9C9; UNIPARC:UPI000000C528; GB:AE005173; NID:g7109463; PDB:1A01
C:Genetics:
A:Gene: F2P9.4
A:Map position: 1
C:Superfamily: alcohol sulfotransferase

Alignment Scores:
Pred. No.: 4, 19e-12 Length: 350
Score: 267.00 Matches: 83
Percent Similarity: 42.6% Conservatives: 56
Best Local Similarity: 25.5% Mismatches: 127
Query Match: 6.1% Indels: 60
DB: 2 Gaps: 12

US-10-768-158-1 (1-2419) x H96768 (1-350)

24	QY	CGCGAGAGCGCGCGAGACCCCGGCGGAGTTTCGAGACGAAGTACTTTC	80
19	Ddb	SerHisAspGluThrLysThrGluSerGluPheGluLysAsnGlnLysArgTyrGln	38
81	QY	-----GAG 83	
39	Ddb	AspLeuIleSerThrPheProHisGluLysGlyTyrArgProLysGluProLeuIleGlu	58
84	QY	TTCCATGCGGTG---CGGCTGCGCGGCGGAGATGCGGAGAGATCGCCAAC	140
59	Ddb	TyrGlyGlyTyrTrpTrpLeuProSerLeuLeuGluGlyCysIleHisalagInGluPhe	78
141	QY	TTCCCGGTGCGGCGCGAGCGATGGATCGTCACTACCCCAAGTCGCGACCACTTG	200
79	Ddb	PheGlnAlaArgProSerAspPheLeuValCysSerTyrProLysThrGlyThrTrp	98
201	QY	CTGCAGGAGGTGCTACTTGTGTAGCCAGGCGGTGACCCCGATGAG-----	248
99	Ddb	LeuTyrAlaLeuThrPheAlaIleAlaAsnArgSerArgPheAspSerSerAsnPro	118
249	QY	ATCGGCTTGATGAACATGACGACGAGCTCCGGTCTCTGAG-----TACCCA	299
119	Ddb	LeuLeuLysArgAsnProHisGluPheValProTyrIleGluIleAspPheProPhe	138
300	QY	CGGCGCTTGGACATCATCAAGCACTGACCTCTCCCGCTCATCAAGACCACTGCC	359
139	Ddb	ProGluValAspValLeuLysAsp---LysGlyAsnThrLeuPheSerThrHisIlePro	157
360	QY	TACCGCTTTGCGCTCTGACCTCCACATGAGACTCCAGGTCACTATATGCTCGC	419
158	Ddb	TyrGluLeuLeuProAspSerValValLysSerGlyCysLysMetValTyrIleTrp	177
420	QY	AACCCCAAGGATCTGTTGTGCTTATATCATAGTTCCACCGCTCTCTCGGGACC	473
178	Ddb	GluProLysAspThrPheIleSerMetTrpThrPheLeuHisLysGluArgThrGluLeu	197
474	QY	-----ATGAGCTACCGGCGACCTTTCAAGAAATTCCTCGCGGAGGTTTATGAT	524
198	Ddb	GlyProValSerAsnLeuGluGluSerPheAspMetPheCysArgGlyLeuSer	215
525	QY	AAGTGGGCTACGGTCTCTGGTTTGAGCAGTGCAGGAGTCTGGGAGCACCGCATG	584
216	Ddb	-----GlyTyrGlyProTyrLeuAsnHisIleLeuAlaTyrTrpLysAlaTyrGlnGlu	233

C;Accession: S69188
R;Jacomme, C.; Roby, D.
Plant Mol. Biol. 30, 995-1008, 1996
A;Title: Molecular cloning of a sulfotransferase in Arabidopsis thaliana and regulation
A;Reference number: S69188; MUID:96270377; PMID:8639757
A;Accession: S69188
A;Molecule type: mRNA
A;Residues: 1-302 <IAC>
A;Cross-references: UNIPARC:UPI0000046BEB; EMBL:Z46823; NID:g599639; PIDN:CRA86850.1; PI
C;Superfamily: alcohol sulfotransferase
C;Keywords: sulfotransferase

Alignment Scores:
Pred. No.: 1.16e-11 Length: 302
Score: 261.00 Matches: 82
Percent Similarity: 44.7% Conservatives: 53
Best Local Similarity: 27.2% Mismatches: 115
Query Match: 5.9% Indels: 52
DB: 2 Gaps: 13

US-10-768-158-1 (1-2419) x S69188 (1-302)

QY 6 GCGGAGCGCGCGCATGGCGAGAGCGGCGGAGAGCCCGCCAGCACCCCG----- 56
Db 12 GlyAspGluAspLeuThrGlnGluThrArgAlaLeuLeuSerSerLeuProLysGluLys 31

QY 57 GCGGAGTTCGAGAGCAAGTACTTCGAGTTCATGGCGTG---CGGCTGCCGCCCTTCTG 113
Db 32 GlyTrpLeuValSerGluIleTyrGluPheGlnGlyLeuTrpHisThrGlnAlaIleLeu 51

QY 114 CCGGGAGAGATGGAGAGATCGCAACTTCCCGGTGCGGCCGCGGAGCGAGTGTGGATCGTC 173
Db 52 GlnGlyIleLeuIleCysGlnLysArgPheGluAlaLysAspSerAspIleIleLeuVal 71

QY 174 ACCTACCCCAAGTCCGCGACACCTGTCGACGAGTGTCTACTTGTG----- 224
Db 72 ThrAsnProLysSerGlyThrTrpLeuLysAlaLeuValPheAlaLeuLeuAsnArg 91

QY 225 -----AGCCAGGCGCTGACCCCGCATGAGATCGGCTTGTATGAACATC 266
Db 92 HisLysPheProValSerSerSerGlyAsnHisPro-----LeuLeuValThrAsnPro 109

QY 267 GAGGACGACCTCCCGTCTCTGGAG-----TACCCACAGCGCGGCTCGACATCATC 317
Db 110 HisLeuLeuValProPheLeuGluGlyValTyrGluSerProAspPheAsp---Phe 128

QY 318 AAGGAATCAGCTCTCCCGCCTCATCAAGAGCCACTGCCCTTCTGCGCCTCT 377
Db 129 SerSerLeuProSerProArgLeuMetAsnThrHisIleSerHisLeuSerLeuProGlu 148

QY 378 GACCTCCACATGGAGACTCCAAAGTCACTATATGGCTCGCAACCCCAAGGATCTGTG 437
Db 149 SerValLysSerSerSerCysLysIleValTyrCysCysArgAsnProLysAspMetPhe 168

QY 438 GTGTCTTATATCAGTTCACCGCTCTCTG-----CGGACATGAGCTACCGAGGC 488
Db 169 ValSerLeuTrpHisPheGlyLysLysLeuAlaProGluGluThrAlaAspTyr----- 186

QY 489 ACCTTTCAAGATTCTCGCGGAGTTTATGATGATAAGCTGGCTACCGCTCTGGTTT 548
Db 187 ProIleGluLysAlaValGluAlaPheCysGluGlyLysPheIleGlyGlyProPheTrp 206

QY 549 GAGCAGCTGAGGAGTTCGG-----GAGCACCGCATGGACTCGAACCTGCTTTTCTC 602
Db 207 AspHisIleLeuGluTyrTrpTyrAlaSerArgGluAsnProAsnLysValLeuPheVal 226

QY 603 AGATATGACATGATCGGACCTGGTGACATGCTGGAGCAGCTGGCCGAGATTCCTG 662
Db 227 ThrTyrGluGluLeuLysLysGlnThrGluValGluMetLysArgIleAlaGluPheLeu 246

QY 663 GGGGTCTCTGTGACAAAGCCAGCTGGAAGCCCTGACCGAGCAGCTGCCACAGCTGGTG 722
Db 247 -----GluCysGlyPheIleGluGluGluValArgGlu-----Ile 259

QY 723 GACCACTGCTGCAACGCTGAGGCC-----CTG 749
Db 260 ValLysLeuCysSerPheGluSerLeuSerAsnLeuGluValAsnLysGluGlyLysLeu 279

QY 750 CCGGTGGGC-----CGGGGAAGAGTGGCGCTGTGGAAGGAC 785
Db 280 ProAsnGlyIleGluThrLysThrPhePheArgLysGlyGluIleGlyTrpArgAsp 299

QY 786 ATCTTC 791
Db 300 SerPhe 301

RESULT 53
JC7921
cytosolic sulfotransferase (EC 2.8.2.-) - zebra fish
C;Species: Brachydanio rerio (zebra fish)
C;Date: 31-Mar-2003 #sequence_revision 31-Mar-2003 #text_change 14-Jul-2003
C;Accession: JC7921
R;Sugahara, T.; Liu, C.; Pai, T.G.; Liu, M.C.
Biochem. Biophys. Res. Commun. 300, 725-730, 2003
A;Title: Molecular cloning, expression, and functional characterization of a novel zebr.
A;Reference number: JC7921; MUID:22395662; PMID:12507510
A;Accession: JC7921
A;Molecule type: mRNA
A;Residues: 1-296 <SUG>
A;Cross-references: UNIPARC:UPI000005F06E; GB:AY180110
C;Comment: This enzyme, which belongs to the cytosolic sulfotransferase gene family tha
s containing hydroxyl or amino group, exhibits the enzymatic activities toward a varie
activity.
C;Keywords: sulfotransferase

Alignment Scores:
Pred. No.: 3.6e-11 Length: 296
Score: 254.50 Matches: 75
Percent Similarity: 47.2% Conservatives: 69
Best Local Similarity: 24.6% Mismatches: 122
Query Match: 5.8% Indels: 39
DB: 2 Gaps: 11

US-10-768-158-1 (1-2419) x JC7921 (1-296)

QY 27 GAGAGCGAGCGCGAGACCCCGCGGGGGAGTTCGAGAGCAAGTACTTCGAGTTC 86
Db 5 LysSerArgMetGluThrAlaAlaLysMetLysAsp---GluAspLysLeuTyrArgArg 23

QY 87 CATGCG-----GTGCGGTGCGCGCCTTCTGCGCGGGAGATGGAGGAG 131
Db 24 AspGlyIleLeuTyrSerThrValLeuSerProProGluThrLeuAspLysLeuLysAsp 43

QY 132 ATFCGCCAACTTCCCGTGGCGCCAGCGACGTGTGGATCGTCACCTACCCCAAGTCCGCG 191
Db 44 Leu-----GlnAlaArgGluAspLeuIleLeuValAlaTyrProLysCysGly 60

QY 192 ACCAGCTTGTGACGAGGTGTC---TACTTGTGAGCCAGGCGCTGACCCCGATGAG 248
Db 61 PheAsnTrpMetValAlaValLeuArgLysIleAsnAlaSerThrGlyLysAspGlu 80

QY 249 ATCGCTGTGATAACATCGACGAGCAGCTCCCGGTCTCGGACTCCACAGCGCGGCGCTG 308
Db 81 -----LysProGluArgProProLeuValGluPheLeuProThrVal 96

QY 309 ---GACATCATCAAGGAATGACCTCTCCCGCCTCATCAAGAGCACCTGCCCTACCGC 365
Db 97 GlnGluGluMetAlaGlnMetProProArgLeuLeuGlyThrHisLeuHisProasp 116

QY 366 TTCTGCCCTCTGACCTCCAAATGGAGACTCCAAAGTCACTATATATGCTCGCAACCCC 425
Db 117 AsnMetProAlaThrPhePheThrLysLysProLysIleLeuValValPheArgAsnPro 136

QY 426 AAGGATCTGGTGTCTTATATATCAGTTCCACCGCTCTCTCGGAGCATTGAGTACCGA 485
Db 137 LysAspThrLeuValSerTyrTyrHisPheMetAsnLysAsnProValLeuProAsnAla 156

```
QY 486 GCACCTTTCAGAAATTCTGCGGAGGTTTATGAATGATAAGCTGGGCTACGGCTCCTGG 545
Db 157 GluSerTrpAspLysPheSerAspPheMetThrGlyAspValSerTrpGlySerTyr 176
QY 546 TTTGAGCAGCTGAGGAGTTCTGGGAGCACCAGCTCG---AACGTGCTTTTCTTC 602
Db 177 PheAspHisAlaLeuAla---TrpGluLysArgIleAspAspProAsnValMetIleVal 195
QY 603 AAGTATGAACAGCATCGGACCTGTTGACGATGTTGGAGCAGCATGGCCAGATTCG 662
Db 196 MetTyrGluAspLeuLysGlnAlaLeuProGluGlyValLysLysIleSerGluPhePhe 215
QY 663 GGGGTGCTCTGTGACAGCCAGCTGGAAGCCCTG----- 698
Db 216 SerLeuProLeuThrAspGluGlnValSerIleAlaGlyGlnSerThrPheSerAla 235
QY 699 -----ACGAGCACTGCCACCACTGGTGGACCACTGGTGCACCGCTGAGGCC 746
Db 236 MetValGluAsnSerGlnLysHisGly-----AsnPheGlySer 249
QY 747 CTGCGCGTGGCGGGAAGAGTTGGCTGTGGAAGACATCTTCACCTCTCCATGAAT 806
Db 250 IlePhePheArgLysGlyGluValGlyAspTrpLysAsnHisPheSerGluAlaGlnSer 269
QY 807 GAGAACTTTGACTGTGTATTAACAGAAAGATGGGAAGTGTGACCTC----- 854
Db 270 LysGlnMetAspGluLeuTyrHisSerLysLeuAlaGlyThrLysLeuAlaAlaArgMet 289
QY 855 ACCTTTGACTTTTAT 869
Db 290 AsnTyrAspLeuTyr 294

RESULT 54
A96769
protein flavonol sulfotransferase F2p9.3 [imported] - Arabidopsis thaliana
C,Species: Arabidopsis thaliana (mouse-ear cress)
C,Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C,Accession: A96769
R,Theologas, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.P.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A,Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luro, J.S.; Maity, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A,Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A,Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A,Reference number: A86141; MUID:21016719; PMID:11130712
A,Accession: A96769
A,Status: preliminary
A,Molecule type: DNA
A,Residues: 1-338 <STO>
A,Cross-references: UNIPROT:Q9C9D0; UNIPARC:UPI00000A6FB0; GB:AE005173; NID:g7109462; P
C,Genetics:
A,Map position: 1
C,Superfamily: alcohol sulfotransferase

Alignment Scores:
Pred. No.: 8.88e-11 Length: 338
Score: 249.50 Matches: 68
Percent Similarity: 46.1% Conservative: 57
Best Local Similarity: 25.1% Mismatches: 99
Query Match: 5.7% Indels: 47
DB: 2 Gaps: 10

US-10-768-158-1 (1-2419) x A96769 (1-338)

QY 138 AACTTCCCGTGGCGCCAGCAGCTGTGGATCGTCACCTACCCCAAGTCCGACACAGC 197
Db 66 HisPheGluAlaArgProThrAspPheLeuValCysSerTyrProLysThrGlyThrThr 85
```

```
QY 198 TTGCTGAGGAGGTGGTCTACTTGTGTAGCCAGGGCGCTGACCCCGATGAGATCGGC--- 254
Db 86 TrpLeuLysAlaLeuThrTyrAlaIleValAsnArgSerArgTyrAspAspAlaAlaAsn 105
QY 255 -----TTGATGAACATCGACAGCAGCTCCCGGTCTCTGGAGTACCCCA----- 296
Db 106 ProLeuLeuLysArgAsnProHisGluPheValProTyrValGluIleAspPheAlaPhe 125
QY 297 CAGCGCGGCTGGAGCATCATCAAGAACTGACCTCTCCCGGCTCATCAAGAGCCACCTG 356
Db 126 TyrProThrValAspValLeuGlnAspArgLysAsnPro---LeuPheSerThrHisIle 144
QY 357 CCCTACCGCTTCTGCGCTCTGACCTCCACATGGAGACTCCCAAGTGCATCTATATGGCT 416
Db 145 ProAsnGlyLeuLeuProAspSerIleValAsnSerGlyCysLysMetValTyrIleTrp 164
QY 417 GCACACCCCAAGGATCTGGTGTCTTATTATTCAGTTTC---CACCGTCTCTCTG----- 467
Db 165 ArgAspProLysAspThrPheIleSerMetTrpThrPheLeuHisLysGluLysSerGln 184
QY 468 -----CGGACCATGAGCTACCGGACACCTTTCAAGAAATTTCTCCGGAGGTTTATGAAT 521
Db 185 GluGlyGlnLeuAlaSerLeuGluAspSerPheAspMetPheCysLysGlyLeuSer--- 203
QY 522 GATAAGCTGGGTACGGCTCTCTGGTTTGAGCAGCTGAGGAGTTCTGGGAGCACCGCATG 581
Db 204 -----ValTyrGlyProTyrLeuAspHisValLeuGlyTyrTrpLysAlaTyrGln 220
QY 582 GACTCG-----AACGTGCTTTTCTCAAGTATGAAGACATGCATCGGGACCTGGTGACG 635
Db 221 GluAsnProAspArgIleLeuPheLeuArgTyrGluThrMetArgAlaAsnProLeuPro 240
QY 636 ATGTGTGAGCAGCTGCGCAGATTTCTGGGGGTGTCC---TGTGACAAGGCCAGCTGGNA 692
Db 241 PheValLysArgLeuAlaGluPheMetGlyTyrGlyPheThrAspGluGluGluAsn 260
QY 693 GCCTGTGAGGAGCACTGCCACCACTGTTGGACCACTGTGCAACGCTGAGGCCCTGCC 752
Db 261 GlyValAlaGluLys-----ValValLysLeuCysSerPheGluThrLeuLys 276
QY 752 ----- 752
Db 277 AsnLeuGluAlaAsnLysGlyAspLysGluArgGluAspArgProAlaValTyrAlaAsn 296
QY 753 -----GTGGCGGGGGAAGAGTTGGCTGTGGAAGGACATCTTCACCGCTCTCCATG 803
Db 297 SerAlaTyrPheArgLysGlyLysValGlyAspTrpAlaAsnTyrLeuThrProGluMet 316
QY 804 AATGAGAGTTTGACTTGTGTATATAACAGAAAG 836
Db 317 AlaAlaArgIleAspGlyLeuValGluGluLys 327

RESULT 55
A84523
probable steroid sulfotransferase [imported] - Arabidopsis thaliana
C,Species: Arabidopsis thaliana (mouse-ear cress)
C,Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C,Accession: A84523
R,Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999
A,Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A,Reference number: A84420; MUID:20083487; PMID:10617197
A,Accession: A84523
A,Status: preliminary
A,Molecule type: DNA
A,Residues: 1-333 <STO>
A,Cross-references: UNIPROT:O82330; UNIPARC:UPI000004850D; GB:AE002093; NID:g3650034; P
C,Genetics:
A,Map position: 2
```


C:Superfamily: alcohol sulfotransferase

Alignment Scores:

Pred. No.: 4.66e-10 Length: 333
Score: 240.00 Matches: 80
Percent Similarity: 44.3% Conservative: 67
Best Local Similarity: 24.1% Mismatches: 113
Query Match: 5.5% Indels: 72
DB: 2 Gaps: 15

US-10-768-158-1 (1-2419) x A84523 (1-333)

```
QY 27 GAGAGCGAGCCGAGACCCCGCGGGGAGTTTCAG-----ACCAAGTAC 77
Db 20 GlnThrIleLeuLeuSerSerLeuProThrGluIleAspTyrLeuGlyAsnLysLeu 39
QY 78 TTCGAGTTCCATGGCGTGGCGCTTCTGCGCGGGAAGATGAGAGATCGCC 137
Db 40 PheAsnTyrGluGlyTyrTrp-----TyrSerGluAspIleLeuGlnSerIlePro 56
QY 138 AAC-----TTCGCGGTGCGGCCCGACGAGCTGTGGATCGTCACTACCCCAAG 185
Db 57 AsnIleHisThrGlyPheGlnProGlnGluThrAspIleIleLeuAlaSerPheTyrLys 76
QY 186 TCCGCGACCAAGCTTCTGAGAGGTGCTACTTGTGTCGAGCCAGGGCGCT----- 236
Db 77 SerGlyThrThrTrpLeuLysAlaLeuThrPheAlaLeuValGlnArgSerLysHisSer 96
QY 237 -----GACCCCGATGAGATCGGCTTGATGAAC 263
Db 97 LeuGluAspHisGlnHisProLeuLeuHisAsnProHisGluIle---ValProAsn 115
QY 264 ATCGACGACGACTCCCGGTCTGGAGTACCCACGCGGCGCTGGACATCATCAAGGAA 323
Db 116 LeuGluLeuAspLeu---TyrLeuLysSerSerLysPro-----AspLeuThrLysPhe 132
QY 324 CTG-----ACCTCTCCCGCTCATCAAGAGCCACTGCCCTACCGCTTCTG 371
Db 133 LeuSerSerSerSerProArgLeuPheSerThrHisMetSerLeuAspProLeu 152
QY 372 CCTCTGACTCCACAATGGAGACTCAAGGTTCATCTATATGCTCGCAACCCCAAGGAT 431
Db 153 GlnValProLeuLysGluAsnLeuLysIleValTyrValCysArgAsnValLysAsp 172
QY 432 CTGGTGTGTTTATATCAGTTCCACCGCTCTCTGCGG-----ACCATG 476
Db 173 ValMetValSerValTyrPheArgGlnSerLysLysIleThrArgAlaGluAspTyr 192
QY 477 AGCTACCGAGGACCTTTCAGAAATCTCCCGAGGTTTATGAATGATAAAGCTGGGCTAC 536
Db 193 SerLeuGluAlaIlePheGluSerPheCys-----AsnGlyValThrLeuHis 208
QY 537 GCCTCTGTTGTAGCAGCTGCAGGAGTTCTGGAGCACCGCATGGAC-----TCGRAAC 590
Db 209 GlyProPheThrAspHisAlaLeuSerTyrTrpArgSerLeuGluAspProLysHis 228
QY 591 GTGCTTTTCTCAAGTATGAACATGCACTCGGGACCTGGTGACCATGTGTGGAGCAGCTG 650
Db 229 PheLeuPheMetArgTyrGluAspLeuLysAlaGluProArgThrGlnValLysArgLeu 248
QY 651 GCCAGATTCTGGGGTGTCTGT-----GACAGGCCCGCAGCTG 689
Db 249 AlaGluPheLeu-----AspCysProPheThrLysGluGluGluAspSerGlySerVal 266
QY 690 GAAGCCCTGACGAGCACTGCCACCACTGGTGGACCACTGTGTGCAACGCTGAGGCCCTG 749
Db 267 AspLysIleLeuGluLeuCys-----SerLeuSerAsnLeuArgSerVal 281
QY 750 CCGGTGGGC-----CGGGA 764
Db 282 GluIleAsnLysThrArgThrSerSerArgValAspPheLysSerTyrPheArgLysGly 301
QY 765 AGAGTTGGGCTGTGGGAAGACATCTTCACCGTCTCCATGAATGAGAAGTTTGACTGGTG 824
```

```
Db 302 GlnValGlyAspTrpLysSerTyrMetThrProGluMetValAspLysIleAspMetIle 321
QY 825 TATAAACACAGAGATGGAAAGTGTGACCTTCACGCTTT 860
Db 322 IleGluGluLysLeuLysGlySerGlyLeuLysPhe 333
RESULT 56
T06012
hypothetical protein T25K17.90 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C>Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jul-2004
C:Accession: T06012
R:Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.;
submitted to the Protein Sequence Database, March 1999
A:Reference number: Z15382
A:Accession: T06012
A:Molecule type: DNA
A:Residues: 1-314 <BEV>
A:Cross-references: UNIPROT:Q9STQ6; UNIPARC:UPI00000489F1; EMBL:AL049171; GSPDB:GN000062
A:Experimental source: cultivar Columbia; BAC clone T25K17
C:Genetics:
A:Gene: ATSP.T25K17.90
A:Map position: 4
A:Introns: 109/3
C:Superfamily: alcohol sulfotransferase
Alignment Scores:
Pred. No.: 1.01e-09 Length: 314
Score: 235.50 Matches: 69
Percent Similarity: 47.0% Conservative: 58
Best Local Similarity: 25.6% Mismatches: 104
Query Match: 5.4% Indels: 39
DB: 2 Gaps: 10
US-10-768-158-1 (1-2419) x T06012 (1-314)
QY 108 TTCCTCCCGCGGAAGATGGAGAGATCGCAACTTCCCGGTGCGGCCGACGACGCGTGG 167
Db 46 PheLeuGlnGlyValLeuAsnPheGlnArgGlyPheLysProGlnAspThrAspIleIle 65
QY 168 ATCGTCACCTACCCCAAGTCCCGCACCGAGCTGTGTCGAGGAGTG---GTCTACTTGGT 224
Db 66 ValAlaSerTyrProLysSerGlyThrLeuTrpLeuLysAlaLeuThrValAlaLeuPhe 85
QY 225 AGCCAGGCGCTGACCCGATGAGATCGCTTGATGAACATCGACGAGCAGCTC----- 278
Db 86 GluArgThrLysAsnProSerHis-----AspAspProMetSerHis 99
QY 279 CCGGTCTCTGGAGTACCCACAGCCGCGCTGGACATCATCAAGAACTGACCTTCCCCC 338
Db 100 ProLeuLeuSerAsnAsnProHisAsnLeu-----LeuSerSerSerProArg 116
QY 339 CTCATCAAGACCACTGCCCTTACCGCTTCTGCCCTCTGACCTCCACATGAGACTCC 398
Db 117 LeuPheSerThrHisThrProPheHisThrLeuGlnValAlaValLysAspSerProCys 136
QY 399 AAGGTCACTATATCGCTCGCAACCCCAAGGATCTGGTGTCTTATTATCAGTTCCAC 458
Db 137 LysValValTyrIleCysArgAspAlaLysAspSerLeuValSer-----ArgTrpHis 154
QY 459 -----CGCTCTCTCGGACCATGAGTACCGAGGCACCTTTTCAAGAAATTCGCGCG 509
Db 155 IleValCysArgSerLeuAsnLysGluGluAspArgThrIleLeuGluSerMetPheGlu 174
QY 510 AGGTTTATGAATGATAGCTGGGTACCGCTCTCGCTTGTGGTGGTGGTGGTGGTGGTGG 569
Db 175 SerPheCysSerGlyValCysLeuPheGlyProPheTrpAspHisIleLeuSerTyrTrp 194
QY 570 GAGCACCCATGGAC-----TCGAACGTGCTTTTCTCAAGTATGAAGACATGATCGG 623
Db 195 LysAlaSerLeuGluLysProLysGlnValLeuPheMetArgTyrAspGluIleLysThr 214
```



```
QY 624 GACCTGGTGACCATGGTGGAGCAGCGTGGCCAGATTCCTCGGGGGTCTCTGTGACAGGCC 683
    |||
Db 215 AspProHisGlyGlnLeuLysLysLeuAlaGluPheLeuGly-----CysProPheSer 232
    |||
QY 684 CAGCTGGAGCCCTCAGCGAGCAGTGGCCACAGCTGGTGGACCATGTGTGC-----AAC 737
    |||
Db 233 LysGluGluGluLysAsnGlySerLeuAsnLysIleLeuGluMetCysSerLeuProAsn 252
    |||
QY 738 GCTGAGGCCCTCCCGTGGCC-----758
    |||
Db 253 LeuSerSerLeuGluValAsnLysThrGlyLysSerIleAsnGlyIleGluLysAsn 272
QY 759 -----CGGGGAAGATTGGGTGGTGGGAAGACATCTTCACCGTCTCCATGATGAG 809
    |||
Db 273 HisPheArgLysGlyIleValGlyAspTrpLysAsnHisLeuThrProGluMetGlySer 292
QY 810 AGTTTGTACTGGTGTATAAACAGAGATG 839
    |||
Db 293 LysIleAspMetIleMetLysGluLysLeu 302

RESULT 57
H84451
probable steroid sulfotransferase [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: H84451
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.J. Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.B.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: H84451
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-331 <STO>
A:Cross-references: UNIPROT:Q9ZPQ6; UNIPARC:UPI00000484E3; GB:AE002093; NID:g4406766; PI
C:Genetics:
A:Gene: At2g03750
A:Map position: 2
C:Superfamily: alcohol sulfotransferase

Alignment Scores:
Pred. No.: 7e-09 Length: 331
Score: 224.50 Matches: 68
Percent Similarity: 44.6% Conservative: 53
Best Local Similarity: 25.1% Mismatches: 117
Query Match: 5.1% Indels: 33
DB: 2 Gaps: 8

US-10-768-158-1 (1-2419) x H84451 (1-331)
QY 138 AACTTCCCGTGGCGCCAGCGACGTGGTGGTCACTACCTACCCCAAGTCCGCGCACCAG 197
    |||
Db 63 HisPheLysProArgAspThrAspIleLeuAlaSerLeuProLysGlyThrThr 82
QY 198 TTGCTGCAGGAGGTGCTACTTGGTGAGCCAG-----GGCGTGAC 239
    |||
Db 83 TrpLeuLysSerLeuIlePheAlaValHisArgLysLysThrArgGlyThrProGln 102
QY 240 CCGCATGAGATCGGCTGTGATGAACATCGACGACGCTCCCGTCTCGAG-----290
    |||
Db 103 ThrHisProLeuLeuGlnAsnProHisAspLeuValProPheLeuGluValGluLeu 122
QY 291 TACCCACAGCCGGCTGGACATCATCAAGGAATGACCTCTCCCGCGCTCATCAAGAG 350
    |||
Db 123 TyrAlaAsnSerGlnIleProAspLeuAlaLysTyrSerSerProMetIlePheSerThr 142
QY 351 CACCTGCCCTACCGCTTTCTGCCCTCTGACCTCCCAATGAGAGACTCCAAGTCACTAT 410
    |||
Db 143 HisMetHisLeuGlnAlaLeu---ArgGluAlaThrThrLysAlaCysLysThrValTyr 161
QY 411 ATGGCTCGCNACCCCAAGGATCTGGTGGTCTTATTATCATGTTCCACCGCTCTCTG---467
```

```
Db 162 ValCysArgGlyIleLysAspThrPheValserGlyTrpHisTyrArgAsnMetLeuHis 181
    |||
QY 468 CGGACCATGAGCTACCGAGGCACCTTTCACAAATTCCTCCGAGGTTTGAATGAATAG 527
    |||
Db 192 ArgThrLysMetAspGlnAlaThrPheGluLeuMetPheAspAlaTyr-CysArgGlyVal 201
    |||
QY 528 CTGGCTACGGCTCTCGTTTGAGCAGCTGAGGAGTTCTGGAGCACCAGCATGGACTCG 587
    |||
Db 202 LeuLeuTyrGlyProTyrTrpGluHisValLeuSerTyrTrpLysGlySerLeuGluAla 221
    |||
QY 588 -----AACGTGCTTTTCTCAAGTATGAACACATGATCGGACCTGCTGACGATGGTG 641
    |||
Db 222 LysGluAsnValPheMetLysTyrGluGluIleLeuGluProArgValGlnVal 241
    |||
QY 642 GAGCAGCTGGCCAGATCTCTGGGGGTGCTGTGACAAAGGCCACAGCTGGAAGCCCTGACG 701
    |||
Db 242 LysArgLeuAlaGluPheLeu-----GluCysProPheThrLysGluGluGluSer 259
    |||
QY 702 GAGCACTGCCACGCTGCTGGACCATGCTGTC-----AACGCTGAGGCCCTGCCCTG 755
    |||
Db 260 GlySerValGluGluIleLeuLysLeuCysSerLeuArgAsnLeuSerAsnLeuGluVal 279
    |||
QY 756 GGC-----CGGGGAAGA 767
    |||
Db 280 AsnLysAsnGlyThrThrArgIleGlyValAspSerGlnValPhePheArgLysGlyGlu 299
QY 768 GTTGGCTGTGGAAGACATCTTACCGTCTCCATGAATGAGAGATTGCTGTGTAT 827
    |||
Db 300 ValGlyAspTrpLysAsnHisLeuThrProGlnMetAlaLysThrPheAspGluIleIle 319
    |||
QY 828 AAACAGAGATCGGAAGTGTGACCTCAGCTT 860
    |||
Db 320 AspTyrArgLeuGlyAspSerGlyLeuIlePhe 330

RESULT 58
QBEB3
BHLF1 protein - human herpesvirus 4 (strain B95-8)
C:Species: human herpesvirus 4, Epstein-Barr virus
C:Date: 25-Feb-1985 #sequence_revision 25-Feb-1985 #text_change 09-Jul-2004
C:Accession: A03742
R:Bankier, A.T.; Deininger, P.L.; Farrell, P.J.; Barrell, B.G.
Mol. Biol. Med. 1, 21-45, 1983
A:Title: Sequence analysis of the 17,166 bp EcoRI fragment C of B95-8 Epstein-Barr virus
A:Reference number: A93065; MUID:85035713; PMID:6094825
A:Accession: A03742
A:Molecule type: DNA
A:Residues: 1-660 <BAN>
A:Cross-references: UNIPROT:P03181; UNIPARC:UPI000000CD00
R:Baer, R.; Bankier, A.T.; Biggin, M.D.; Deininger, P.L.; Farrell, P.J.; Gibson, T.J.;
Nature 310, 207-211, 1984
A:Title: DNA sequence and expression of the B95-8 Epstein-Barr virus genome.
A:Reference number: A03794; MUID:84270667; PMID:6087149
A:Contents: annotation; protein coding region
C:Comment: The sequence contains four perfect repeats (residues 149-273, 274-398, 399-573)
C:Superfamily: human herpesvirus 4 BHLF1 protein

Alignment Scores:
Pred. No.: 3.11e-06 Length: 660
Score: 190.50 Matches: 114
Percent Similarity: 32.3% Conservative: 26
Best Local Similarity: 26.3% Mismatches: 152
Query Match: 4.3% Indels: 141
DB: 1 Gaps: 23

US-10-768-158-1 (1-2419) x QBEB3 (1-660)
QY 1079 GGAGAGCTGCAGCTTCTAAAGCGGAGACAGCT-----GCTTCGGTTGGGAATCAT 1029
    |||
Db 173 GlyAlaGlyGlnArgProSerGlyProThrGlyArgProAlaIleProGlyAlaPro 192
    |||
QY 1028 CACACTCCCTCCGCT-----CACGGCGCTCTTCCCTTCCCGCTGTTC 984
    |||
```

Db 193 GlyThrProAlaAlaProGlyProGlyGlyGlyAlaAlaVal-ProSerGlyAlaThrPr 212
QY 983 ACA-
|||
Db 212 oHisProGluArgGlySerGlyProAlaAspProProAlaAlaAlaArgLeuProProGl 232
QY 979 -----GCTGCTCCAGAGTTGTGCCAGAGGAA 951
|||
Db 232 uArgGlnGluProArgLeuProGlnAspLeuAlaAlaGlnArgCysProAlaGlyPr 252
QY 950 TAAATGAATGCATACAGGACTTTTGGCTAGTACTGTCTGGGTATTGTGAGCATGCAGG 891
|||
Db 252 oProProThrArgSerGlyAlaAlaAlaGlnArgThr-----HisArgArgProProGl 270
QY 890 TTGTTTCTCTGTTATTAATAAAGTCAACGGTAGGTCACTTTCCCATCTTCTG 831
|||
Db 270 yCys-----Pro-ArgSerA 275
QY 830 TTTTATACCAAGTCAAACTTCTCATTCATGAGAGCGGTGAAGATGCTCTCCACAGCCC 771
|||
Db 275 laArgAnProGly-----CysProArgThrTrpArgArg----- 286
QY 770 AACTCTTCCCGGCCACGGGCGCTCAGCGTTGCGACACTGGTCCACACAGCTGGTG 711
|||
Db 287 -----ArgSerGlyAlaGlnArg--GlyHisProProProGlyAlaG 300
QY 710 GCAGTCTCGTCAGCGCTTCCAGCTGGGCTTGTACAGGACACCCCGGAGGAATCTGGC 651
|||
Db 300 lGlnArgProSerGly-----ProThrGlyGlyArgP 311
QY 650 CAGCTCTCCA-----CCATCGTCACAGGTCCGGATGCATGCTTCATA 606
|||
Db 311 roAlaAlaProGlyAlaProGlyThrProAlaAlaProGlyProGlyGly----- 327
QY 605 CTTGAGAAAAGCACGTTTCAGTCCATGGCTGCCAGAACTCTCCAGACGTCCTGACGTCCTCAA 546
|||
Db 328 -----GlyAlaAlaValProSerGlyAlaThrProHisProGluArgGlySerG 344
QY 545 CCAGGAGCGGTACCCAGCTTATTCATTCAAAACCTCCGCGAGAATT----- 499
|||
Db 344 lProAlaAspProProAla-----AlaAlaArgLeuProProGluArgG 359
QY 498 -----CTTGAAAGGTCCCTCGGTAGCTCATGGT 471
Db 359 lGluProArgLeuProGlnAspLeuAlaAlaGlnArgCysProAlaGlyProProP 379
QY 470 CCGCAGAGAGCGGTGGAAGTATAAAGACACCCAGATCCTTCGGGTTCG---GAGC 414
Db 379 roThrArgSerGlyAlaAlaAlaGlnArgThrHisArgArgProProGlyCysProArgS 399
QY 413 CATATAGATGACCT-----TGGAGTCTCCATTTGTGAGGTCAGAGGGCAG 369
Db 399 erAlaArgAsnProGlyCysProArgThrTrpArgArgArgSerGlyAlaGlnArgGlyH 419
QY 368 AA-----ACCGTAGGGCAGGTGGCTTGTATGAGAGCGGGGAGAGTCAAGTTC 321
Db 419 isProProGlyAlaGlyGlnArgProSerGlyProThrGlyGlyArgProAlaAlaP 439
QY 320 CTTGATGATGTCAGGC---CCGGCTGTGGTACTCCAGAGCCGGGAGCTGCTCGTCA 264
Db 439 ro---GlyAlaProGlyThrProAla-----AlaProGlyProGlyGlyAlaAlaV 456
QY 263 GTTCATCAACCGCATCTCATCGGGTCAGCGGCTGGCTCACCAGTAGACCACTTCCTG 204
Db 456 aProSerGlyAlaThrProHisProGluArgGlySerGlyProAlaAspProProAlaA 476
QY 203 CAGCAAGCTGGTCCCGGACTTGGGGTAGGTGACGATCCACGTCGCTCGGCGCA---- 148
Db 476 laAlaArgLeuProProGluArgGlnGluProArgLeuProGlnAspLeuAlaAlaAlaG 496
QY 147 -----CCGGGAAGTTGGCGATCTCTCCATCTTCCCGCGGACAGAGGGCGGACCG 96
Db 496 lArgCysProAla-----GlyProProProThrArgSerGlyAlaAlaAlaGlnA 513

QY 95 CAGCCGCTCGAACTCGAAGTACTTGCTCTCGAACTCCCCGGGGTCTCGGGGTCTCGGC 36
Db 513 rgThrHis-----ArgArgProProGlyCysProArgSer---A 525
QY 35 CTCGCTCTCGCATCGCCGCGCGCTCGCGCTCGC 1
Db 525 laArgAnProGlyCysProArgThrTrpArgArg 536
RESULT 59
A95936
probable alcohol sulfotransferase (EC 2.8.2.2) [imported] - Sinorhizobium meliloti (str
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C;Accession: A95936
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Herna
proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A;Title: The complete sequence of the 1.683-kb pSymB megaplasmid from the N2-fixing end
A;Reference number: A95842; MUID:21396508; PMID:11481431
A;Accession: A95936
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-280 <KUR>
A;Cross-references: UNIPROT:Q92VF3; UNIPARC:UPI00000CB647; GB:AL591985; PIDN:CAC49153.1
A;Experimental sources: strain 1021, megaplasmid pSymB
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
neault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.;
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: Smb21249
A;Genome: plasmid
C;Keywords: sulfotransferase
Alignment Scores:
Pred. No.: 0.00449 Length: 280
Score: 161.00 Matches: 64
Percent Similarity: 41.0% Conservative: 46
Best Local Similarity: 23.9% Mismatches: 98
Query Match: 3.7% Indels: 60
DB: 12 Gaps: 12
US-10-768-158-1 (1-2419) x A95936 (1-280)
QY 156 AGCGACGTGTGGATCGTCACTACCCCAAGTCCGCGCACCGAGTTGTCGAGGAGTGGTC 215
Db 25 AlaAspSerPheLeuIleSerTyrProLysSerGlyArgThrTrpPheArgTyrValLeu 44
QY 216 -----TACTTGTGAGCCAGCGCGCTGACCCCGATGAGATCGGCTTGATGAACTCGAC 269
Db 45 SerHisThrLeuAlaThrIleAlaArgValProGluThrIleAspLeuHisAsnMetPhe 64
QY 270 GAGCAGTCCCGGTCTGGAGTACCCACAGCGCGGCTGACATCATCAAGGAACGTGACC 329
Db 65 SerIleValProAsnPheAsp-----LeuAspProValArgGlyMetPro 79
QY 330 TCT-----CCCGGCTCATCAAGACCCACTTGGCCC 359
Db 80 GlyTyrArgPheArgGluAlaLysAspAlaIleProThrIleLeuValSerHisLeuAsp 99
QY 360 TACCGC-----TTCTGCGCTCTGACCTCCCAATGGAGACTCCCAAGGTCACTAT 410
Db 100 TyrArgAlaSerLeuPheLeu-----ArgArgProValIleMet 112
QY 411 ATGGCTCCCAACCCCAAGGATCTGGTGTCTTATTATCATGTTCCACCGCTCTCGCG 470
Db 113 MetValArgAspProArgAspValIleValSerAlaTyr---PheHisAlaThrArgHis 131

```
QY 471 ACCATGAGCTACCGAGGACACCTTTCAAGAAATCTCCCGAGGTTTATGAATCATTAAGCTG 530
Db   ::: |||||
132 LysHisArgPheAlaGlyThrLeuThrGluPheIleLys-----AspArgAspGln 148
QY 531 GCCTACGGCTCTCGTTTGAGCACCTGCGAGGAGTTCTCGGAGCACCGCATCGGATCGAAC 590
Db   ::::::::::|
149 GlyMetProLysMetIleAspTyrLeuAsnGly---TpAlaAlaGlyLeuSerAsnArg 167
QY 591 GTGCTTTT---CTCAAGATGAAACATGCAATCGGACCTGGTGACGATGGTGAGCAG 647
Db   |||||
168 AlaHisPheValLeuSerTyrGluGlyLeuSerAlaAspThrAspGlyArgThrGluAla 187
QY 648 CTGGCCAGATTCTCGGGGTCTCTGTGACAGGCCCGACCTGGAAGCCCTGACGGAG--- 704
Db   ::::::::::|
188 ValLeuGluPheLeuArgCysProValAspArgThrAlaLeuArgAlaAlaValGluAla 207
QY 705 -----CACTGCCAC 713
Db   |||||
208 GlyArgPheGluAlaMetGlnAspArgGluArgValGluGlyIleProAlaHisAspTyr 227
QY 714 CAGCTGGTGACACGTGTCGAACCGCTGAGGCCCTCGCGTGGCGGGGAGAGTTGGG 773
Db   |||||
228 AspArgAsnAsp-----ValGluSerLeuArgMetArgArgGlyLysAlaGly 243
QY 774 CTGCTGGAAGGACATCTTACCCTCTCCATGAATGAGAAGTTTGACTTGGTGATAAACAG 833
Db   ::::::::::|
244 GlyPheArgAspTyrLeu-----AspGluAlaGlnValAlaGluValGlu 258
QY 834 AAGATGGGAAGTGTGACCTCACG 857
Db   |||||
259 ArgLeuCysAlaAlaAspLeuThr 266
```

RESULT 60

A45344

immediate-early protein - suid herpesvirus 1 (strain Kaplan)

C:Species: suid herpesvirus 1

C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004

C:Accession: A45344

R:Vlcek, C.; Kozmik, Z.; Paces, V.; Schirm, S.; Schwyzler, M.

Virology 179, 365-377, 1990

A:Title: Pseudorabies virus immediate-early gene overlaps with an oppositely oriented op

A:Reference number: A45344; PMID:91021039; PMID:2171211

A:Accession: A45344

A>Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-1446 <VLC>

A:Cross-references: UNIPROT:P33479; UNIPARC:UPI000012D219; GB:M34651; NID:g334070; PIDN:

C:Superfamily: herpesvirus immediate-early protein IE175

C:Keywords: DNA binding; early protein; transcription regulation

Alignment Scores:

Pred. No.:	0.00109	Length:	1446
Score:	158.00	Matches:	96
Percent Similarity:	32.1%	Conservative:	19
Best Local Similarity:	26.8%	Mismatches:	102
Query Match:	3.6%	Indels:	141
DB:	1	Gaps:	16

US-10-768-158-1 (1-2419) x A45344 (1-1446)

```
QY 759 GGCACACGGGCGGCGCTCAGCGTTGCAGCACTGGTCCA-----CCAGCTGGT 712
Db   |||||:::
136 GlyProArgSer-----ArgAlaGlySerGlyProArgProThrProAlaAla 152
QY 711 ---GGCAGTGTCCGTCCAGGCTTCCAGCTGGGCGCTTGTCCACAGACACCCCGAGAAATC 655
Db   |||||
153 LeuAlaAlaGluAlaGlyAlaProGlyGlyProGlyArgSerSerProSerAlaAla 172
QY 654 TGGCCAGCTGCTCCACCATGTCACCGAGTCCCGATGTCATGTCTTCATCTTGAGAAAAA 595
Db   |||||
173 SerProAlaSerSerGlySerProGlyProSerAlaAlaProArgArgTrpSerPro 192
QY 594 GCACGTTCCAGTCCATGCGGT-----GCTCCAGAACTCTCTGCA--- 556
```

RESULT 61

CGCH15

collagen alpha 1(I) chain - chicken (tentative sequence) (fragments)

C:Species: gallus gallus (chicken)

C>Date: 12-Aug-1981 #sequence_revision 06-Jul-1982 #text_change 31-Mar-2000

C:Accession: A90458; A90181; A02857

R:Highberger, J.H.; Corbett, C.; Dixit, S.N.; Yu, W.; Seyer, J.M.; Kang, A.H.; Gross, J

Biochemistry 21, 2048-2055, 1982

A:Title: Amino acid sequence of chick skin collagen alpha1(I)-C8 and the complete prim-

A:Reference number: A90458; PMID:82231995; PMID:7093229

A:Accession: A90458

A:Molecule type: protein

A:Residues: 1-1036 <HIG>

A:Cross-references: UNIPARC:UPI0000173B62

A;Experimental source: skin
A;Note: this is the latest in a series of papers from these workers elucidating the sequence of the alpha 1(I) chain of type I procollagen.
R;Eyre, D.R.; Glimcher, M.J.
Biochem. Biophys. Res. Commun. 48, 720-726, 1972
A;Title: Evidence for a previously undetected sequence at the carboxyterminus of the alpha 1(I) chain of type I procollagen.
A;Reference number: A90181; MUID:72243016; PMID:5047697
A;Accession: A90181
A;Molecule type: protein
A;Residues: 1037-1042 <EYR>
A;Cross-references: UNIPARC:UPI0000173863
A;Experimental source: skin
A;Note: residues 1037-1042 above correspond to the carboxyl end of the protein
C;Comment: Lysines at positions 103, 700, 934, and 946 above may be hydroxylated in some species.
C;Comment: Most of the prolines at the third position of the tripeptide repeating unit (Pro-Gly-Hyp) are hydroxylated.
C;Comment: Pro-1002 is the only 3-hydroxyproline and the only hydroxylated proline in the alpha 1(I) chain.
C;Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology; C;Keywords: coiled coil; extracellular matrix; glycoprotein; pyroglutamic acid; trimer; F;1/Modified site: pyroglutamic acid (Gln) #status experimental

Alignment Scores:
Pred. No.: 0.0011 Length: 1042
Score: 157.50 Matches: 81
Percent Similarity: 28.7% Conservative: 10
Best Local Similarity: 25.6% Mismatches: 121
Query Match: 3.6% Indels: 105
DB: 1 Gaps: 16

US-10-768-158-1 (1-2419) x CGCH1S (1-1042)

Qy	783	CTTCTCCACAGCCCAACTCTTCCCGGCCCA-----CGGCAGGGCCTCAGCGTTGCAGC	730
Db	708	ProProGlyAlaThrGlyPheProGlyAlaAlaGlyArgValGlyProProGlyProSer	727
Qy	729	-----ACTGGTCCACCACT-----	715
Db	728	GlyAsnIleGlyLeuProGlyProProGlyProAlaGlyLys**GlySerLysGlyPro	747
Qy	714	---GGTGGCAGTGCTCGTCAGGCTTCCAGCTGGGCTTGTCCACAGGACACCCACGGA	658
Db	748	ArgGlyGluThrGlyProAlaGlyArgProGlyGluProGlyProAlaGlyProProGly	767
Qy	657	ATCTGGCCAGCTGTCTCCACCATCTCCACAGTCCCGATCGATCTCTTCATCTACTTGAA	598
Db	768	-----ProProGlyGluLysGlySerProGlyAlaasp-----	778
Qy	597	AAAGCAGTTCGAGTCCATCGGTCTCCAGAACTCTCGACGTCTCAACACGAGAGC	538
Db	779	-----GlyProIleGlyAlaProGlyThrProGlyProGlnGlyIleAlaGly	794
Qy	537	CGTAGCCAGCTTATCATTAACCTCCGGCAGATTCTTGAAGTGCCTC-----	484
Db	795	GlnArgGlyValValGly-----LeuProGlyGlnArgGlyGluArgGlyPheProGly	812
Qy	483	-----GCTAGCTCATGG-----TCCGACAGAGCGGTGGAACTGATAATAAG	442
Db	813	LeuProGlyProSerGlyGluProGlyLysGlnGlyProSerGlyAlaSerGlyGluArg	832
Qy	441	ACACACACAGATCTCTGGGTTCGAGCCATATAGATGACTTGGAGTCTCCATTGTGGA	382
Db	833	GlyProProGlyProMetGly-----ProProGlyLeu-----Ala	844
Qy	381	GGTCAGAGGGCAAGCGGTAGGAGGTGGCTCTTGATGAGCGGGGAGAGGTCAATT	322
Db	845	GlyProProGlyGluAlaGlyArgGluGly-----	854
Qy	321	CTTGTGATGATCTCCAGGCCCGCTGTGGGTACTCCAGGACCGGAGCTGCTCGTCGATGT	262
Db	855	-----AlaProGly-----AlaGluGlyAlaProGlyArgaspGlyAlaAlaGlyPro	870
Qy	261	TCATCAGCCGATCTATCGGGTTCAG-----CGCCCTCGCTCACCAGT	217
Db	871	LysGlyAspArgGlyGluThrGlyProAlaGlyProProGlyAlaProGlyAlaProGly	890


```
Db 236 ProProGly---LyProGlnGlyProProGlnGln-----GlyGlyAsnArgPro 251
Qy 369 GAAAGCGTAGGCGAGGTGGCTCTTGATGATGAGCGGGAGAGTCAAGTCTTCCTTGATGATGT 310
Db 252 GlnGlyProProPro----- 256
Qy 309 CCAGCCCGGCTGTGGTACTCCAGGACGGGAGCTGCTCGATGTTTCATCAAGCCGA 250
Db 257 ProGlyLyProGlnGlyProProProGlnGlyAspLySerArgSerProGlnSer--- 275
Qy 249 TCTCATCGGGTCAGCGCCCTCGCTCACCAGTAGACACCTCTCCAGCAGCAAGCTGGTGC 190
Db 276 -----ProProGlyLyProGlnGlyProProProGln----- 286
Qy 189 CGGACTTGGGTAGTGCATCCACATCCTCCTCGGCGCACCGGGAAGTTGGCGATCT 130
Db 287 -----GlyGlyAsnGlnProGlnGlyProProProProGlyLyProGlnGly 303
Qy 129 CTTCATCTCCCGCGGCGAGAGG-----GCG 103
Db 304 ProProProGlnGlyGlyAsnLyProGlnGlyProProProProGlyLyProGlnGly 323
Qy 102 GCAGCGCGCAGCATGGAAGTCTGCTCTCGAAGTCTTCTCGAAGTCTCCCGCGGGTCTGGGGG 43
Db 324 ProProAlaGlnGlyGlySerLySerGlnSerAlaArgAlaProProGly----- 340
Qy 42 TCTCGGCTCTCGTCTCCGCCA 22
Db 341 ---LyProGlnGlyProPro 346
```

RESULT 65

```
B40505
Hypothetical protein - suid herpesvirus 1 (strain Indiana-Funkhuser or Becker)
C:Species: suid herpesvirus 1
C>Date: 10-Apr-1992 #sequence_revision 10-Apr-1992 #text_change 09-Jul-2004
C:Accession: B40505
R:Cheung, A.K.
J. Virol. 65, 5260-5271, 1991
A:Title: Cloning of the latency gene and the early protein 0 gene of pseudorabies virus.
A:Reference number: A40505; MUID:91374576; PMID:165441
A:Accession: B40505
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1958 <CH>
A:Cross-references: UNIPROT:Q69340; UNIPARC:UPI00000F2308; GB:M57505; NID:g334066; PIDN:
C:Superfamily: pseudorabies virus 1 nuclear antigen
```

Alignment Scores:

Pred. No.:	0.00214	Length:	1958
Score:	154.50	Matches:	83
Percent Similarity:	31.5%	Conservative:	18
Best Local Similarity:	25.9%	Mismatches:	112
Query Match:	3.5%	Indels:	108
DB:	2	Gaps:	15

US-10-768-158-1 (1-2419) x B40505 (1-1958)

```
Qy 849 CACACTTCCCATCTTCTGTTATACACCAAGTCAAACTTCTCAT----- 805
Db 317 HisGlnAspProProGlyProProThrProThrProSerThrSerHisHisHisGln 336
Qy 804 -----TCATGGAGACGGTGAAGATGCTTCCACAGCCCAACTCTTCCCGGC 757
Db 337 GlyProProThrSerProArg-----ProSerThrSerHisGlnAspPro 352
Qy 756 CCACGGCGCGGCTCAGCGTTGCAGCACT-----GTCACCAAGCT 715
Db 353 ProGlyGlyGlyProProSerAlaGluThrHisHisHisHisGlnAspProProGly 372
Qy 714 GTGGCAGTGTCTCGTCAGGGCTCCAGCTGGCGCTTGCACAGGACACCCCGAAGTC 655
Db 373 GlyGly-----ProProSerThrSerSerHisHisHisGlnAspProProGlyGly 390
```

```
Qy 654 TGGCAGCTGCTCCACCATCGTCCACAGTCCCGATCGATGCTCTTATATCTTGAGAAAAA 595
Db 391 GlyProProSerProProProArgProSerThrSerSerSer-----SerSer 406
Qy 594 GCAGCTTCGAGTCCATCGGTGCTCCAGAACTCTCGACCTGCTGCTCAACACGAGGACCGT 535
Db 407 HisGlnGlyProProSerThrArgProProProGlnArgProProArgProPro 426
Qy 534 AGCCAGCTTATCATTTATAAACCTCCGGCAGATTCTTGAAGGTGCTCGGTAGCTCA 475
Db 427 ProPro-----SerProGlnLySerGluThrArgAlaGlySerGlu 441
Qy 474 TGGTCCGAGAGAGG-----GGTGAAGTATATAAGACACACACAGATCTTGGGGT 421
Db 442 AsnThrAlaGlnThrLeuPheSerHisSerGluAsnLyLeuPheSerHisProMetGly 461
Qy 420 TCGAGCCATATAGATGACCTTGGAGTCTCCATTGTGGAGTCCAGAGGCGAAGACCGGT 361
Db 462 -----GluGlyGlyGluGlyAspArgGly 469
Qy 360 AGGCAGGTGCTCTTGATGAGGCGGGGAGAGGTCA-----GTTCTTGA 316
Db 470 ThrAlaGlyGlyGlu-----GlyAspArgAspAspProArgProProSerProPro 486
Qy 315 TGATGTCCAGCGCGGCTGCGGTACTCCAGGACCGGAGCTGCTCGTATGTTTCATCA 256
Db 487 ProArgProProProProProProProProProPro----- 498
Qy 255 AGCCGATCTCATCGGGGTGAGCGGCTCGCTCACCAGTAGACACCACTCTTCAGCAAGC 196
Db 499 -----ProProProProGlnProProProAlaGly--- 509
Qy 195 TGGTCCCGAGCTTGGGTAGTGAGTCCACAGTCCAGCTGCTGCGCGCACCGGAGATTGG 136
Db 510 -----GlySerAlaArgArgArgArg-----GlyGlyGly 520
Qy 135 CGATCTCTCCATCTTCCCGGAGAGGCGGCGGCGCACGCG-----CATGGA 85
Db 521 GlyGlyProProGlyArgGlyGlyArgArgGlyGlyLeuArgArgAlaGluGly 540
Qy 84 ACTCGAAGTACTTGTCTCTC----- 66
Db 541 Thr-GluAlaAlaAlaAlaAspAlaGluGluGluAspGlyAspGlyAspGluAspGlu 560
Qy 65 -----GAATCTCCCGCGGCTGCTGGGG 43
Db 560 uAspArgAlaGluAspGlyArgGluAspGlyGlyGlyGlyProArgGlyAlaGlyGly 580
Qy 42 T 42
Db 580 Y 580
```

RESULT 66

WMBE38

```
infected cell protein ICP34.5 - human herpesvirus 1 (strain F)
C:Species: human herpesvirus 1
C>Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 09-Jul-2004
C:Accession: A27768
R:Chou, J.; Roizman, B.
J. Virol. 57, 629-637, 1986
A:Title: The terminal a sequence of the herpes simplex virus genome contains the promoter.
A:Reference number: A27768; MUID:86115412; PMID:3003394
A:Accession: A27768
A:Molecule type: DNA
A:Residues: 1-358 <CHO>
A:Cross-references: UNIPROT:P08353; UNIPARC:UPI0000170E8C; GB:M12240; NID:g330122; PIDN:
C:Comment: This protein accumulates late in infection predominantly in the cytoplasm of
C:Genetics:
A:Gene: ICP34.5
C:Superfamily: herpesvirus infected cell protein ICP34.5
F:Keywords: tandem repeat
F:175-204/Region: 3-residue repeats (A-T-P)
```


QY 171 CGATCCACAGTCGTGGCGCCACCGGAGAGT-----TGGCGATCTCCTCCACTTCC 118
 Db 379 -----PrometProGlySerAlaProArgProGlyProProAlaSer 392
 QY 117 CGCGCAGAGGCGCGCAGCCAGCCATCGAAGTCTGCTCGAAGTCC 58
 Db 393 AlaAlaAlaSerGlyProAlaArgProArgAlaAlaValAlaProCybValArgAlaPro 412
 QY 57 CGCGGTCTGGGGTCTCGGCTCGCTCTCCG 25
 Db 413 ProPro-----GlyProGlyProArgAlaPro 421
 RESULT 68
 S21626
 collagen alpha 1(I) chain precursor - mouse
 C:Species: Mus musculus (house mouse)
 C>Date: 13-Jan-1995 #sequence revision 25-Apr-1997 #text change 09-Jul-2004
 C:Accession: S57243; S16374; A23982; I49559; I49557; S39789; I48300; S21626
 R:Li, S.W.; Khillan, J.; Prockop, D.J.
 Matrix Biol. 14, 593-595, 1994
 A:Title: The complete cDNA coding sequence for the mouse pro-alpha-1(I) chain of type I
 A:Reference number: S57243
 A:Accession: S57243
 A:Molecule type: mRNA
 A:Residues: 1-1453 <LIS>
 A:Cross-references: UNIPROT:P11087; UNIPARC:UPI0000027558; EMBL:U08020; NID:g470673; PID
 R:Metzger, M.; Toman, D.; De Crombrughe, B.; Vuorio, E.
 Biochim. Biophys. Acta 1089, 241-243, 1991
 A:Title: Specific hybridization probes for mouse type I, II, III and IX collagen mRNAs.
 A:Reference number: S16176; MUID:91274355; PMID:2054384
 A:Accession: S16374
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1442-1453 <MET>
 A:Cross-references: UNIPARC:UPI0000000691; EMBL:X57981; NID:g50484; PIDN:CAA41046.1; PID
 R:French, B.T.; Lee, W.H.; Maul, G.G.
 Gene 39, 311-312, 1985
 A:Title: Nucleotide sequence of a cDNA clone for mouse proalpha1(I) collagen protein.
 A:Reference number: A23982; MUID:86137403; PMID:3841523
 A:Accession: A23982
 A:Molecule type: mRNA
 A:Residues: 518-1128 <PRE>
 A:Cross-references: UNIPARC:UPI000016CC46; GB:M14423; NID:g192261; PIDN:AAA37333.1; PID:
 R:Monson, J.M.; Friedman, J.; McCarthy, B.J.
 Mol. Cell. Biol. 2, 1362-1371, 1982
 A:Title: DNA sequence analysis of a mouse pro-alpha-1(I) procollagen gene: Evidence for
 A:Reference number: I49559; MUID:83141374; PMID:6298597
 A:Accession: I49559
 A>Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 735-1130 <RES>
 A:Cross-references: UNIPARC:UPI000016CC47; GB:M17491; NID:g192263; PIDN:AAA37334.1; PID:
 R:Harbers, K.; Kuehn, M.; Delius, H.; Jaenisch, R.
 Proc. Natl. Acad. Sci. U.S.A. 81, 1504-1508, 1984
 A:Title: Insertion of retrovirus into the first intron of alpha1(I) collagen gene leads
 A:Reference number: I49557; MUID:84170331; PMID:6324198
 A:Accession: I49557
 A>Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-25 <RE2>
 A:Cross-references: UNIPARC:UPI000005B79; GB:K01688; NID:g192246; PIDN:AAA37330.1; PID:
 R:Fenton, S.P.; Laman, S.R.; Hannagan, M.; Stacey, A.; Jaenisch, R.; Bateman, J.F.
 Biochim. Biophys. Acta 1216, 469-474, 1993
 A:Title: Genomic sequence of mouse COL1A1 encoding the collagen propeptides.
 A:Reference number: S39789; MUID:94092741; PMID:8268229
 A:Accession: S39789
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-185, 1203-1201, 'G', 1203-1218, 'E', 1220-1221, 'T', 1223-1
 A:Cross-references: UNIPARC:UPI000005B65; UNIPARC:UPI000017738A
 R:Rhodes, K.; Rippe, R.A.; Umezawa, A.; Nehls, M.; Brenner, D.A.; Breindl, M.
 Mol. Cell. Biol. 14, 5950-5960, 1994
 A:Title: DNA methylation represses the murine alpha 1(I) collagen promoter by an indirect
 A:Reference number: I48300; MUID:94344105; PMID:8065328

A:Accession: I48300
 A>Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: UNIPARC:UPI000017738B; EMBL:X54876; NID:g50486; PIDN:CAA38657.1; PID
 C:Genetics:
 A:Gene: COL1A1
 A:Introns: 770/3; 788/3; 806/3; 842/3; 860/3; 932/3; 968/3; 1004/3; 1023/3; 1058/3; 1058/3
 C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;
 C:Keywords: coiled coil; extracellular matrix; glycoprotein; heterotrimer; triple helix
 F:1-22/Domain: signal sequence #status predicted <SIG>
 F:23-151/Domain: amino-terminal propeptide #status predicted <PRO>
 F:30-89/Domain: von Willebrand factor type C repeat homology <WVC>
 F:152-1453/Product: collagen alpha 1(I) chain #status predicted <MAT>
 F:1224-1453/Domain: fibrillar collagen carboxyl-terminal homology <FCC>
 Alignment Scores:
 Pred. No.: 0.00239 Length: 1453
 Score: 153.50 Matches: 84
 Percent Similarity: 28.9% Conservative: 8
 Best Local Similarity: 26.4% Mismatches: 119
 Query Match: 3.5% Indels: 107
 DB: 2 Gaps: 18
 US-10-768-158-1 (1-2419) x S21626 (1-1453)
 QY 783 CTTCCACAGCCCAACTCTTCCCGGCCCA-----CGGCAGGCGCTCAGGTTGCAGC 730
 Db 859 ProProGlyAlaThrGlyPheProGlyAlaAlaGlyArgGlyProProGlyProSer 878
 QY 729 -----ACTGGTCCACCACT-----GGTGGCAGTCTCCG 700
 Db 879 GlyAsnAlaGlyProProGlyProProGlyProValGlyLysGluGlyLysGlyPro 898
 QY 699 TCAGGCTTCCAGCTGGCGCTTGTCCAGGACA-----CCCCA 661
 Db 899 ArgGly---GluThrGlyProAlaGlyArgProGlyGluValGlyProProGlyProPro 917
 QY 660 GGAATCTGGCAGCTGCTCCACCATCGTCACAGTCCCGATGTCATGCTTCATCTGA 601
 Db 918 Gly-----ProAlaGlyGluLysGlySerProGlyAlaAsp----- 929
 QY 600 GAAAAAGCAGCTCGAGTCCATCGCGTCTCCAGAACTCTTCGACGTGCTCAAAACAGG 541
 Db 930 -----GlyProAlaGlySerProGlyThrProGlyProGlnGlyLeuAla 944
 QY 540 AGCGTAGCCACAGCTATCATTAACCTCCGCGAGAAATCTTTGAAGGTGCCTC--- 484
 Db 945 GlyGlnArgGlyValValGly-----LeuProGlyGlnArgGlyGluArgGlyPhePro 962
 QY 483 -----GGTAGCTCATGG-----TCCGACAGAGCGGTGGAAGTATAAT 445
 Db 963 GlyLeuProGlyProSerGlyGluProGlyLysGlnGlyProSerGlySerGlyGlu 982
 QY 444 AAGACACACACAGATCTTGGGGTTGCGAGCCATATAGATACCTTGGAGTCTCCATTGT 385
 Db 983 ArgGlyProProGlyProMetGly-----ProProGlyLeu----- 994
 QY 384 GGAGTTCAGAGGCGAGAAAGCGTAGGCGAGTGGCTCTTGTATAGGCGGGAGAGGTCA 325
 Db 995 AlaGlyProProGlyGluSerGlyArgGluGlySer----- 1006
 QY 324 GTTCTTGTATGATGTCAGGCCCGCTGTGGGTACTCCAGGACCGGAGCTGCTCGTCA 265
 Db 1007 -----ProGly---AlaGluGlySerProGlyArgAspGlyAlaProGly 1020
 QY 264 TGTTCATCAAGCCGATCTCATCGGGTTCAG-----CGCCCTGGCTCACCA 220
 Db 1021 AlaLysGlyAspArgGlyGluThrGlyProAlaGlyProProGlyAlaProGlyAlaPro 1040
 QY 219 AGTAGACCACTCTTCGACCAAGCTGGTCCCGAGCTTTGGGTAGGTGACGATCCACCT 160
 Db 1041 GlyAlaProGlyProValGlyProAlaGlyLys-----AsnGlyAspArg 1055

QY	159	CGTGGCCGCCACCGGGAAGTTGGCGANTCTCTCCATCTTCCCGCGGACGAAGGCGGCA	100
Db	1056	GlyGluThrGlyProAlaGlyProAlaGlyProIleGlyProAlaGlyAlaArgGlyPro	1075
QY	99	GCCGCGCCGATCGAACTCGAAGTACTTGCTCTCGAACT	61
Db	1076	AlaGlyProGlnGlyProArgGlyAspLysGlyGluThrGlyGluGlnGlyAspArgGly	1095
QY	60	-----CCCCCGGGGTGCTGGGGTCT---	40
Db	1096	IleLysGlyHisArgGlyPheSerGlyLeuGlnGlyProProGlySerProGlySerPro	1115
QY	39	-----CGGCTCGCTCTCGGCA	22
Db	1116	GlyGluGlnGlyProSerGlyAlaSerGlyProAlaGlyProArgGlyProPro	1133
RESULT 69			
S23809			
collagen alpha 2(I) chain homolog - sea urchin (Strongylocentrotus purpuratus)			
C:Species: Strongylocentrotus purpuratus (purple urchin)			
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004			
A:Accession: S23809			
R:Exposito, J.Y.; d'Alessio, M.; Solursh, M.; Ramirez, F.			
J. Biol. Chem. 267, 15559-15562, 1992			
A:Title: Sea urchin collagen evolutionarily homologous to vertebrate pro-alpha-2(I) coll			
A:Reference number: S23809; MUID:92348411; PMID:1639795			
A:Status: preliminary			
A:Molecule type: mRNA			
A:Residues: 1-1414 <EXP>			
A:Cross-references: UNIPROT:Q26634; UNIPARC:UPI000007BD61; EMBL:M92040; NID:g161435; PID			
C:Superfamily: collagen alpha 2(I) chain; fibrillar collagen carboxyl-terminal homology			
C:Keywords: coiled coil; extracellular matrix; glycoprotein; trimer; triple helix			
F:1207-1414/Domain: fibrillar collagen carboxyl-terminal homology <FCC>			
Alignment Scores:			
Pred. No.:	0.00259	Length:	1414
Score:	153.00	Matches:	77
Percent Similarity:	32.9%	Conservative:	19
Best Local Similarity:	26.4%	Mismatches:	102
Query Match:	3.5%	Indels:	94
DB:	1	Gaps:	17
US-10-768-158-1 (1-2419) x S23809 (1-1414)			
QY	762	CCGCGCCACGGCGGCTCAGCGTTCAGCACTGGTCCACAGCTGGTGGCGAGTGTCT	703
Db	66	ProGlyProValGlyProGlyProGlyProSerGlyProSerGlyProAlaGlyAsnAsnGly	85
QY	702	CGTCAGGCTTCCAGCTGGGCTTGTTCACAGACACCCCGGAAATC-----	655
Db	86	ProProGly--ProAsnGlyPro-----ArgGlyAsnProGlyMetAspGlyLeuThr	102
QY	654	-----TGGCGAGTGTCTCACCATCG-----	634
Db	103	GlyLeuProGlyIleProGlyProProGlyProProGlyLysSerGlySerLeuValAla	122
QY	633	-----TCACCAGTCCCGATGCATGCTTTCATACATTGAGAAAAA	595
Db	123	SerAlaGlnThrSerSerPheAsnLysGlyProSerLeuAlaGlyTyrGlnTyrProGln	142
QY	594	GCACGTTCGAGTCCATCGCGTGTCTCCAGAACTCTCTGCACGTGCTCAACACGAGGACCGT	535
Db	143	AlaGlnAlaAlaGlyThrProGlyProArgGlyProPro--GlyProProGlySerArg	161
QY	534	ACCCAGCTTATCATTAACCTCCGCGAGAAATCTTTGAAAGGTGCTCGGTACTCA	475
Db	162	GlyProGlnGlyLeuThrGlyProSerGlyPro-----	172
QY	474	TGTTCCGACAGAGCGGTGGAAGTATGATATAAGACACACACAGATCTTGGGTTCCGAG	415
Db	173	----SerGlyGluThrGlyProSerGlyAsnSerGlyProProGlyProSerGlyLeu---	190

QY	414	CCATATAGATGACCTTGGAGTCTCCATTGTGAGGTTCAGAGGCGCAGAAAGCCGTAGGCA	355
Db	191	ProGlyArg-----ProGlySerAspGlyAspGlyThrPro	203
QY	354	GTTGGCTCTTGATGAGCGGGGAGAGGTCAAGTCTTGTGATGATCCAGGCGCGGTGTG	295
Db	204	GlySerGln-----GlyGlnArg-----GlyProAla---	212
QY	294	GGTACTCCAGGA-----CCGGAGCTGCTCGTCCA-----	265
Db	213	GlyThrProGlySerArgGlyThrProGlyMetProGlyAlaProGlyMetLysGlyHis	232
QY	264	-----TGTTTCATCAAGCCGATCTCATCGGGGTGAG-----CGCCCTGGC	226
Db	233	GlnGlyLeuProGlyMetThrGlySerLysGlyGluArgGlyGluGlyGluArgGly	252
QY	225	TCACCAAGTAGACCACTCTCGCAGCAAGCTGGTCCCGACTTGGGTAGGTGACGATCC	166
Db	253	SerAspGlySerProGlyProValGlyAla-----	262
QY	165	ACACGTGCTGGGCGGCACCGGGAAGTTGGCGATCTCTCCATCTTCCCGCGCAGAAGG	106
Db	263	-----ProGlyProAlaGlyProSerGlyGlnProGlyGluArgGlyArgThr	278
QY	105	GCGGAGCGGCACGCACTGAACCTGA-----AGTACTTGCTCTCGAACTCCC	58
Db	279	GlyProAlaGlySerGlnGlyAspArgGlyAlaAspGlyAlaThrGlySerGlnGlyPro	298
QY	57	CCGGGGTGTGGGGGTCTCGGCTCGCTCTCCGCCA	22
Db	299	ProGlySerThrGly-----ProAlaGlyAlaPro	308
RESULT 70			
S13141			
hypothetical protein (ribosomal RNA repeat region) - Giardia lamblia			
C:Species: Giardia lamblia			
C:Date: 06-Dec-1996 #sequence_revision 06-Dec-1996 #text_change 05-Oct-2004			
A:Accession: S13141; S10886			
R:Upcroft, J.A.; Healey, A.; Mitchell, R.; Boreham, P.F.L.; Upcroft, P.			
Nucleic Acids Res. 18, 7077-7081, 1990			
A:Title: Antigen expression from the ribosomal DNA repeat unit of Giardia intestinalis.			
A:Reference number: S13141; MUID:91088287; PMID:2263466			
A:Accession: S13141			
A:Molecule type: DNA			
A:Residues: 1-294 <UPC>			
A:Cross-references: UNIPROT:O9XZV7; UNIPARC:UPI0000177CC5; EMBL:X52949			
A>Note: the source is designated as Giardia intestinalis			
A>Note: readthrough of the terminator TAG is supposed to occur between residues 241-Ala			
R:Healey, A.; Mitchell, R.; Upcroft, J.A.; Boreham, P.F.L.; Upcroft, P.			
Nucleic Acids Res. 18, 4006, 1990			
A:Title: Complete nucleotide sequence of the ribosomal RNA tandem repeat unit from Giar			
A:Reference number: S10886; MUID:90326542; PMID:2374731			
A:Accession: S10886			
A:Status: translation not shown			
A:Molecule type: DNA			
A:Residues: 1-241 <HEA>			
A:Cross-references: UNIPARC:UPI0000177CC6; EMBL:X52949			
A>Note: the source is designated as Giardia intestinalis			
A>Note: the assignment of the coding region has been revised in reference S13141			
C:Superfamily: Proline-rich peptide P-B			
Alignment Scores:			
Pred. No.:	0.00219	Length:	294
Score:	152.00	Matches:	83
Percent Similarity:	28.8%	Conservative:	10
Best Local Similarity:	25.7%	Mismatches:	98
Query Match:	3.5%	Indels:	132
DB:	2	Gaps:	18
US-10-768-158-1 (1-2419) x S13141 (1-294)			
QY	759	GGCCACCGGCGGCGCTCAGCGTTGACGACTGTGCCACGCTGGT-----	712

QY 504 AGAATCTTGAAGGTGCTCGGTAGCTCATGGTCGCAGAGAGCGGTGGAAGTGAAT 445
 Db LysGlyAspArgGlyGluAsnGlySer-----ProGlyAlaProGlyHis 1047
 QY 444 AAGACACACACAGATCTTGGGGTTGCGAGCCATATAGATGACCTTGGAGTCTCCATTGT 385
 Db ProGlyProProGlyProValGlyProSerGlyLys----- 1059
 QY 384 CGAGGTTCAGAGCGCAAGAGCGGTAGGCGAGGTGGCTCTTGTAGAGCGGGAGAGGTCA 325
 Db SerGlyAspArgGlyGluThrGlyProAlaGly----- 1070
 QY 324 GTTCTTGTATGATGCCAGGCCG-----GCTGTGGGTACTCCAGGACCGGAGCTGCT 271
 Db ProSerGlyAlaProGlyProAlaGlyAlaArgGlyAlaProGlyProGlnGlyPro 1089
 QY 270 CQTCGATGTTTCATCAAGCCGATCTCATCGGGTTCAGCGC----- 232
 Db ArgGlyAspLysGlyGluThrGlyGluArgGlySerAsnGlyLysGlyHisArgGly 1109
 QY 231 ---CCTGGCTCACCAGTAGACCA---CCTCTGCAGCAAGCTGGTGGCGACTTGGGT 178
 Db PheProGlyAsnProGlyProProGlySerProGlyAla-AlaGlyHisGlnGlyAla 1129
 QY 177 AGGTACCATCCACAGCTCGCTGGCGCGCACCGGGAAGTTGGCGATCTCTCCATCTCC 118
 Db SerProGlyProAlaGlyProArgGlyProValGlyPro-----HisGlyPro 1146
 QY 117 CGCGCAGAGGGCGGAGCGCACCGCCATCGAAGTACTGCTCTCGAAGTCCC 58
 Db ProGly-LysAspGlyThr-----SerGlyHisP 1156
 QY 57 CGCGGTCTGGGGTCTCGGCTCGCTCTCGCCATCGCGCGCGTGGCGG 5
 Db roGlyProIleGly-----ProProGlyPro 1164
 RESULT 72
 70345
 hypothetical protein KIAA0324 - human (fragment)
 C:Species: Homo sapiens (man)
 C>Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 09-Jul-2004
 C:Accession: T02345
 R:Ricker, D.O.; Bruce, D.; Mundt, M.; Doggett, N.; Munk, C.; Saunders, E.; Robinson, D.;
 re, J.; White, S.; Ueng, S.; Tatum, O.; Campbell, C.; Fawcett, J.; Deaven, L.
 submitted to the EMBL Data Library, March 1998
 A:Description: Sequencing of human chromosome 16p13.3.
 A:Reference number: Z14664
 A:Accession: T02345
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1791 <RIC>
 A:Cross-references: UNIPROT:O60382; UNIPARC:UPI00000730AE; EMBL:AC004493; NID:g2996648;
 C:Genetics:
 A:Map position: 16
 A:Introns: 1610/2; 1706/2
 A:Note: KIAA0324
 Alignment Scores:
 Pred. No.: 0.00355 Length: 1791
 Score: 151.50 Matches: 98
 Percent Similarity: 34.6% Conservative: 22
 Best Local Similarity: 28.2% Mismatches: 121
 Query Match: 3.4% Indels: 106
 DB: 2 Gaps: 14
 US-10-768-158-1 (1-2419) x T02345 (1-1791)

QY 794 GGTGAAGATGCTTCCACAGCCCAACTCTTCCCGCGCCAGCGGCGCTCAGCGTT 735
 Db LysGlyAspArgGlyGluAsnGlySer-----ProGlyAlaProGlyHis 1047
 QY 649 GlySerAspSerProGlyProLysAla-ProAlaProArgAlaLeuProArgArGs 668
 QY 734 GCAGACTGGTCCACACAGCTGGTGCAGTCTCGTTCAGGGCTTCACGCTGGCGCTTGT 675
 Db LysGlyAspArgGlyGluThrGlyProAlaGly----- 1070

RESULT 73

A54849

collagen alpha 1(VII) chain precursor - human
 N:Alternate names: procollagen alpha 1(VII) chain
 C:Species: Homo sapiens (man)
 C>Date: 04-Nov-1994 #sequence_revision 04-Nov-1994 #text_change 09-Jul-2004
 C:Accession: A54849; P08044; S16316; 156328; A30296; I84686

Db 668 erArgSerGlySerSerSerLysGlyArgGlyProSer-----ProGluGlySerSers 686
 QY 674 ACAGACACACCCCGAGAAATCTGGCCAGCTGCTCCACCA----- 637
 Db 686 erThrGluSerSer-----ProGluHisProProLysSerArgThrAlaArgG 703
 QY 636 -----TCGTACACAGGTCCCGATGATGCTTTCATCTACTTGAAGAAAGCAGCTTCA 585
 Db 703 LysArgSerSerProGluProLysThrLysSerArgThrProArgArgArGsSers 723
 QY 584 GTCCATCGGTGCTCCAGAACTCTCGCACGTCT----- 550
 Db 723 erArg-----SerSerProGluLeuThrArgLysAlaArgLysSerArgSerArgSera 742
 QY 549 -----CAAAACACAGGAGCCGTAGCCACAGCTTATCAT-----TCAT 516
 Db 742 LaSerSerSerProGluThrArgSerArgThrProProArgHisArgArgSerProSerV 762
 QY 515 AAACCTCGGAGAAATCTTGAAGGTGCTCGGTAGTCTATGTCCTCCGAGAGCGGTG 456
 Db 762 alSerSerProGluProAlaGluLysSerArgSerArgArgArgSerAlaSerS 782
 QY 455 GAACTGATAATAAGACACACCATCTCTGGGGTTGCGAGCCATATAGATGACCTTGA 396
 Db 782 erProArgThrLys--ThrThrSerArgGly-----Ar 793
 QY 395 GTCTCCA-----TTGTGAGGTTCAGAGGCGAGGAGCGTAGGCGAG 354
 Db 793 gSerProSerProLysProArgGlyLeuGlnArgSerArgSerArgSerArgGluL 813
 QY 353 GTGGCTCTTGATGAGCGGAGAGGTTCAGTCTCTTGATGATGCTC----- 308
 Db 813 sThrArgThrThrArgArgArgAspArgSerGlySerGlnSerThrSerArgArg 833
 QY 307 -----AGGCCGCGCTGGGTGCTAC----- 290
 Db 833 gLlnArgSerArgSerArgSerArgValThrArgArgArgGlyGlySerGlyThi 853
 QY 289 -----TCCAGGACCGGAGAGCTGCTCGTCTGATGTTTCAT 258
 Db 853 sSerArgSerProAlaArgGlnGluSerSerArgThrSerSer-ArgArgArgGlyA 873
 QY 257 CAAGCCGATCTATCGGGGTGAGCGCTGCTGCTACCAAGTAGACACCTCTCTGAGCAA 198
 Db 873 rgSerArgThrProProThrSerArgLysArgSerArgSerArgThrSerProAla----- 891
 QY 197 GCTGTGCGGACCTTGGGTAGTGCATCCACAGTCCGCTGGCGCGCAGGGAAGTT 138
 Db 892 -----ProThrLysArgSerArgSerArg-----AlaSerProAlaThrH 905
 QY 137 GCGCATCTCTCCATCT-----TCCCGCGCGCAGGAGG 105
 Db 905 isArgArgSerArgSerArgThrProLeuLleSerArgArgSerArgSerArgThrS 925
 QY 104 CGGAGCGCGCAGCGCATGGAACCTGCACTGCTCTCGAACT----- 61
 Db 925 erProValSerArgArgSerArgSerArgThrSerValThrArgArgSerArgS 945
 QY 60 -----CCCCCGGGTCTGGGGTCTCGGCTCTCGCTCTCCGCGCATCCCGCGCC 12
 Db 945 erArgAlaSerProValSerArgArgSerArgSerArgThrProProValThrArgA 965
 QY 11 GTCGCGGTGCG 1
 Db 965 rGAArgSerArg 968

R:Christiano, A.M.; Greenspan, D.S.; Lee, S.; Uitto, J.
 J. Biol. Chem. 269, 20256-20262, 1994
 A:Title: Cloning of human type VII collagen. Complete primary sequence of the alpha1(VII)
 A:Reference number: A54849; MUID:94327588; PMID:8051117
 A:Accession: A54849
 A>Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-2944 <CHR>
 A:Cross-references: UNIPROT:Q02388; UNIPARC:UPI000017A138; GB:L02870; NID:g987124; PIDN:
 R:Tanaka, T.; Takahashi, K.; Furukawa, F.; Inamura, S.
 Biochem. Biophys. Res. Commun. 183, 958-963, 1992
 A:Title: Molecular cloning and characterization of type VII collagen cDNA.
 A:Reference number: PH0844; MUID:92231302; PMID:1567409
 A:Accession: PH0844
 A:Molecule type: mRNA
 A:Residues: 'EPR', 340-475, 'RALSTASHSTLCWRATRWHPNCRGSHWTRAAACEPCNRPAASHRAARAG', 524-528, 'C',
 A:Cross-references: UNIPARC:UPI000017A139; DDBJ:D11152; DDBJ:D13694; NID:g453698; PIDN:E
 A:Experimental source: keratinocyte
 A:Note: The authors translated the codon ACC for residues 394 and 397 as Tyr
 R:Parente, M.G.; Chung, L.C.; Rynnaenen, J.; Woodley, D.T.; Wynn, K.C.; Bauer, E.A.; Mac
 Proc. Natl. Acad. Sci. U.S.A. 88, 6931-6935, 1991
 A:Title: Human type VII collagen: cDNA cloning and chromosomal mapping of the gene.
 A:Reference number: S16316; MUID:91334380; PMID:1871109
 A:Accession: S16316
 A:Molecule type: mRNA
 A:Residues: 815-892, 'E', 894-1439 <PAR>
 A:Cross-references: UNIPARC:UPI000016A722; GB:M65158; GB:S49017; NID:g180914; PIDN:AAA96
 A:Experimental source: keratinocyte
 R:Gannon, W.R.; Abernethy, M.L.; Padilla, K.M.; Priasayanh, P.S.; Cook, M.E.; Wright, J.;
 J. Invest. Dermatol. 99, 691-696, 1992
 A:Title: Noncollagenous (NC1) domain of collagen VII resembles multidomain adhesion prot
 A:Reference number: 156328; MUID:93107742; PMID:1469284
 A:Accession: 156328
 A>Status: translated from GB/EMBL/DDBJ
 A:Molecule type: mRNA
 A:Residues: 'EPR', 372-517, 'DV', 520-540, 'W', 542-1255 <RES>
 A:Cross-references: UNIPARC:UPI000016B3AC; GB:S51236; NID:g262308; PIDN:AAB24637.1; PID:
 R:Settler, J.L.; Eissen, A.Z.; Bauer, E.A.; Morris, N.P.; Glanville, R.W.; Burgeson, R.E.
 J. Biol. Chem. 264, 3822-3826, 1989
 A:Title: Cleavage of type VII collagen by interstitial collagenase and type IV collagena
 A:Reference number: A30296; MUID:89139437; PMID:2537292
 A:Accession: A30296
 A:Molecule type: protein
 A:Residues: 'A', 1240-1246, 'G', 1248-1250, 'XE', 1253-1255, 'Q', 1257, 'E', 2032, 'C', 2034-2041;
 A:Cross-references: UNIPARC:UPI000017A13A; UNIPARC:UPI000017A13B; UNIPARC:UPI000017A13C;
 A:Note: two reported peptides cannot be reliably located
 R:Greenspan, D.S.
 Hum. Mol. Genet. 2, 273-278, 1993
 A:Title: The carboxyl-terminal half of type VII collagen, including the non-collagenous
 A:Reference number: 184686
 A:Accession: 184686
 A>Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: mRNA
 A:Residues: 2395-2871, 'S', 2873-2944 <RES>
 A:Cross-references: UNIPARC:UPI000016A724; GB:L06862; NID:g388713; PIDN:AAA89196.1; PID:
 R:Christiano, A.M.; Rynnaenen, M.; Uitto, J.
 Proc. Natl. Acad. Sci. U.S.A. 91, 3549-3553, 1994
 A:Title: Dominant dystrophic epidermolysis bullosa: identification of a Gly --> Ser subs
 A:Reference number: A55255; MUID:94224777; PMID:8170945
 A:Content: annotation
 C:Comment: Prolines and lysines at the third position of the tripeptide repeating unit
 ed and subsequently O-glycosylated.
 C:Genetics:
 A:Gene: GDB:COL7A1; EBR1; EBD1; EB
 A:Cross-references: GDB:128750; OMIM:120120
 A:Map position: 3p21.3-3p21.3
 A:Note: defects in this gene can result in dominant and recessive dystrophic epidermolys
 A:Note: there are 118 introns
 C:Complex: type VII collagen is probably a homotrimer
 C:Function:
 A:Description: structural component of extracellular polymer associated with anchoring f
 C:Keywords: coiled coil; extracellular matrix; glycoprotein; hydroxylysine; hydroxyproli
 F:1-16/Domain: signal sequence #status predicted <SIG>

F:17-2944/Product: collagen alpha 1(VII) chain #status predicted <MAT>
 F:17-1253/Domain: amino-terminal nonhelical #status predicted <NCL>
 F:136-201/Domain: von Willebrand factor type A repeat homology <VWA1>
 F:231-318/Domain: fibronectin type III repeat homology <FN1>
 F:327-413/Domain: fibronectin type III repeat homology <FN2>
 F:414-502/Domain: fibronectin type III repeat homology <FN3>
 F:508-593/Domain: fibronectin type III repeat homology <FN4>
 F:598-683/Domain: fibronectin type III repeat homology <FN5>
 F:686-771/Domain: fibronectin type III repeat homology <FN6>
 F:776-862/Domain: fibronectin type III repeat homology <FN7>
 F:864-952/Domain: fibronectin type III repeat homology <FN8>
 F:954-1045/Domain: fibronectin type III repeat homology <FN9>
 F:1052-1219/Domain: von Willebrand factor type A repeat homology <VWA2>
 F:1170-1172/Region: cell attachment (R-G-D) motif
 F:1189-1253/Region: cysteine/proline-rich
 F:1254-2783/Region: interrupted helical
 F:1334-1336/Region: cell attachment (R-G-D) motif
 F:2008-2010/Region: cell attachment (R-G-D) motif
 F:2553-2555/Region: cell attachment (R-G-D) motif
 F:2784-2944/Domain: carboxyl-terminal nonhelical #status predicted <NC2>
 F:2876-2929/Domain: animal Kunitz-type proteinase inhibitor homology <BPI>
 F:337-786,1109/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:2167,2176,2185,2188,2664,2667,2673/Modified site: 4-hydroxyproline (Pro) #status exper
 F:2625,2631/Modified site: 5-hydroxylysine (Lys) #status experimental
 F:2625,2631/Binding site: carbohydrate (Lys) (covalent) #status experimental
 F:2634,2802,2804/Disulfide bonds: interchain #status predicted

Alignment Scores:

Pred. No.:	0.00431	Length:	2944
Score:	151.00	Matches:	83
Percent Similarity:	33.8%	Conservative:	15
Best Local Similarity:	28.6%	Mismatches:	126
Query Match:	3.4%	Indels:	66
DB:	2	Gaps:	15

US-10-768-158-1 (1-2419) x A54849 (1-2944)

QY	762	CCCCGCCACGGGCGGCGCTTCCAGCTTGCAGCTTGCAGCTTCCACGAGCTGGTGGCAGTGCT	703
DB	2185	ProGlyProProGlyAlaPro-----GlyLeuAlaGlyProAlaGlyProGlnGly	2201
QY	702	CCGTCAGGGCTTCCAGCTGGGCGCTTGTCCACAGGACACCCAGGAATCTGCCAGCTGCT	643
DB	2202	ProSerGlyLeuLysGlyGluProGlyGluThrGlyProGlyArgGlyLeuThrGly	2221
QY	642	CCA-----CCATCGTCACAGTCCCGATGCGATGCTTCATCACTGA	601
DB	2222	ProThrGlyAlaValGlyLeuProGlyProGlyProSerGlyLeuVal-----	2238
QY	600	GAATAAGACAGTTCAGTCCATGCGGTCTCCAGAACTCTCCAGACGT---GCTCAAAAC	544
DB	2239	-----GlyProGlnGlySerProGlyLeuProGlyGlnValGlyGluThr	2253
QY	543	AGGAGCCGTAGCCAGCTTATCATTAACCTCCGCGAGAAATCTTTGAAAGTGCCTC	484
DB	2254	GlyLysProGlyAlaProGlyArgAspGlyAlaSerGlyLysAspGlyArg-----	2271
QY	483	GGTAGCTCATGTCGCGAGAGCGGTGGAAGTGAATGATTAACACACCACCATCTTGG	424
DB	2272	GlySerProGlyValProGlySerProGlyLeu-----ProGlyProVal	2286
QY	423	GGTTCGAGCCATATAGATGAGCTTGGAGTCTCCATTTGGAGGTTCAGAGGCGAGAAAC	364
DB	2287	Gly-----ProGlyGlyGluProGlyPheThrGlyAlaProGlyGln---AlaValVal	2303
QY	363	GGTAGGCGAGGTGGCTTCTTGATGAGCGGGGAGAGTCACTTCTTGATGTGCAGGC	304
DB	2304	GlyLeuProGlyAlaLys-----GlyGluLysGlyAlaProGlyGlyLeuAlaGly	2320
QY	303	CCGGCTGGGTACTCCAGGACCGGAGCTGCTCGATGTTTCATCAAGCCGATCTCAT	244
DB	2321	AspLeuValGlyGluProGlyAlaLysGlyAspArgGlyLeuProGlyProArgGlyGlu	2340

QY 243 CGGGTCAG-----CGCCCTGGCTCACCAGTAGACACCTCCCTGCACCAAGCTGGTCC 190
Db :|||:||||:|||||
2341 LysGlyGluAlaGlyArgAlaGlyGluProGlyAspProGlyGluGly-GlnLysGlyAl 2360
QY 189 C-----GGAGTTGGGTAGGTGACGATCCACA 163
Db :|||:||||:|||||
2360 aProGlyProLysGlyPheLysGlyAspAspProGlyValGlyValProGlySerProGl 2380
QY 162 CTTGCTGGCGCCGACCGGGAAGTTGGCATCTCTCCATC----- 122
Db :|||:||||:|||||
2380 yProProGlyProProGlyValLysGlyAspLeu-GlyLeuProGlyLeuProGlyAlap 2400
QY 121 -----TTCCCGCGCAGAGGCG-----GGCAGCCGACGC 91
Db :|||:||||:|||||
2400 roGlyValValGlyPheProGlyGlnThrGlyProArgGlyGluMetGlyGlnProGlyP 2420
QY 90 CATGGAACTCGAAGTACTGCTCTCGAACTCCCGGGGTGCTGGGGTCTCGGCTCGC 31
Db :|||:||||:|||||
2420 ro---SerGlyGluArgGlyLeuAlaGlyProProGlyArgGluGlyLleProGlyProL 2439
QY 30 TCTCCGCCATCGCGCGCGCTCGCGC 5
Db :|||:||||:|||||
2439 euGly-----ProProGlyPro 2444

RESULT 74
149607
procollagen type V alpha 2 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C:Accession: I49607
R.Andrikopoulos, K.; Suzuki, H.R.; Solursh, M.; Ramirez, F.
Dev. Dyn. 195, 113-120, 1992
A:Title: Localization of pro-alpha 2(V) collagen transcripts in the tissues of the devel
A:Reference number: I49607; PMID:93214071; PMID:1297453
A:Accession: I49607
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-1497 <RES>
A:Cross-references: UNIPROT:Q61431; UNIPARC:UPI0000028657; GB:I02918; NID:g309180; PIDN:
C:Genetics:
A:Gene: Col5a-2
C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;
F:39-98/Domain: von Willebrand factor type C repeat homology <WMC>
F:1270-1497/Domain: fibrillar collagen carboxyl-terminal homology <FCC>

Alignment Scores:
Pred. No.: 0.00406 Length: 1497
Score: 150.50 Matches: 83
Percent Similarity: 30.1% Conservative: 11
Best Local Similarity: 26.6% Mismatches: 123
Query Match: 3.4% Indels: 95
DB: 2 Gaps: 16

US-10-768-158-1 (1-2419) x I49607 (1-1497)

QY 783 CTTTCCACAGCCCAACTCTCCCGGCCCA-----CGGGCAGGCGCTCAGCGTTGCAGC 730
Db :|||:||||:|||||
902 ProProGlyAlaThrGlyPheProGlySerAlaGlyArgValGlyProPro----- 918
QY 729 ACTGGTCCACAGCTGGTGGCAGTCTCGTCAGGCTTCCAGCTGGG---CCTGTGCAC 673
Db :|||:||||:|||||
919 -----GlyProAlaGlyAlaProGlyProAlaGly---ProAlaGlyGluProGlyLys 935
QY 672 AGGACACACCCAGGAATC-----TGG 652
Db :|||:||||:|||||
936 GluGlyProProGlyLeuArgGlyAspProGlySerHisGlyArgValGlyAspArgGly 955
QY 651 CCAGCTGCTCCACCATCGTCACCAAGTCCCGATGCTTTTCATCTTGAGAAAAGCA 592
Db :|||:||||:|||||
956 ProAlaGlyProProGlySerProGlyAspLysGly-----AspProGlyGluAspGly 973
QY 591 CTTTCAGTCCATGGTCTCTCCAGAACTCTCTGCAGCTGCTCAAAACAGGAGC----- 538
Db :|||:||||:|||||

Db 974 GlnProGlyProAspGlyProProGlyProAlaGlyThrThrGlyGlnArgGlyIleVal 993
QY 537 -----CGTAGCCCACTTATCAT 520
Db 994 GlyMetProGlyGlnArgGlyValThrGlyMetProGlyLeuProGlyProAlaGly--- 1012
QY 519 TCATAAACTCCGCGCAGAAATCTTGAAGGTGCCTCGGTAGCTCATGTCGCCACAGAGC 460
Db :|||:||||:|||||
1013 -----ThrProGlyLysVal-----GlyProThr 1020
QY 459 GGTGAACCTGATAAAGACACACCAGATCCTTGGGTGTGCGAG-----CCA 412
Db 1021 GlyAlaThrGlyAspLysGlyProProGlyProAlaGlyProProGlySerAsnGlyPro 1040
QY 411 TATAGACACCTTGGAGTCTCATTTGCGAGTCCAGAGGGCAGAAAGCGGTAGGACAGT 352
Db 1041 ValGlyGluProGlyProGluGlyProAlaGlyAsnAspGlyThrProGlyArgAspGly 1060
QY 351 GGCTCTTGATGAGCGGGGAGAGGTTCAGTTCCTTGATGATGTCACGGCCCG----- 301
Db 1061 AlaValGlyGluArgGlyAspArgGly-----AspProGlyProAlaGlyLeu 1076
QY 300 -----GCTGTGGTACTCCAGGACCGGAGCTGCTCGTATTCATCA 256
Db 1077 ProGlySerGlnGlyAlaProGlyThrProGlyProValGlyAlaProGlyAspAlaGly 1096
QY 255 AGCCGATCTCATCGGGGTGAGCGCTCCACCAAGTAGACCACTCTCTCCACCAAGC 196
Db 1097 GlnArgGlyGluProGlySerArg-----GlyProValGlyProProGlyArgAla--- 1113
QY 195 TGGTCCCGACTTGGGGTAGGTGACGATCCACACGTCGCTGGCGCCGACCGGGAAGTTGG 136
Db 1114 -----GlyLysArgGlyLeuProGlyProGlnGlyPro 1124
QY 135 CGATCTCTCCATCTTCCCGGCGGAGAGGGCGGCGCAGCGCATGTGAACTCGAAGT 76
Db 1125 ArgGlyAspLysGlyAspAsnGlyAspArgGlyAspArgGlyGlnLysGlyHisArgGly 1144
QY 75 ACTTGCTCTCGAACT-----CCCGGGGTGCTGGGGTCTCGG----- 37
Db 1145 PheThrGlyLeuGlnGlyLeuProGlyProGlyProAsnGlyGluGlnGlySerAla 1164
QY 36 -----CCTCGCTCTCGGCCA 22
Db 1165 GlyIleProGlyProPheGlyProArgGlyProPro 1176

RESULT 75
T45025
mucin MUC5B, tracheobronchial [imported] - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
C:Accession: T45025
R.Desseyn, J.L.; Guyonnet-Duperat, V.; Porchet, N.; Aubert, J.P.; Laine, A.
J. Biol. Chem. 272, 3168-3178, 1997
A:Title: Human mucin gene MUC5B, the 10.7 kb large central exon encodes various alternat
A:Reference number: 222899; MUID:97166151; PMID:9013550
A:Accession: T45025
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-3570 <DES>
A:Cross-references: UNIPARC:UPI0000016AD88; EMBL:Z72496; NID:g1834502; PIDN:CAA96577.1; P
A:Experimental source: placenta
C:Genetics:
A:Gene: MUC5B

Alignment Scores:
Pred. No.: 0.00491 Length: 3570
Score: 150.50 Matches: 206
Percent Similarity: 29.8% Conservative: 83
Best Local Similarity: 21.2% Mismatches: 326
Query Match: 3.4% Indels: 357
DB: 2 Gaps: 42

Db	152	Val	-----GlyProAlaGlyProAlaGlyAspGlnGlyArgHis	164
QY	441	ACACCACCATCTTGGGTGGAG	-----CCATATAGATGACCTTGGAGTCTCCAT	388
Db	165	GlyProProGlyProThrGlyGlyGlnGlyPro	: :	176
QY	387	TGTGGAGGTGAGAGGGCAGAAAGCGGTAGGCAGGTGGCTTTGATGAGGCGGGGAGAGG	: :	328
Db	177	-----GlyGluGlnGlyAspAlaGlyArgProGlyAlaAla	: :	188
QY	327	TCAGTTCTTCATGATGATCCAGGCCCGGCTGTGGGTACTCCAGGACCGGAGCTGCTCGT	: :	268
Db	189	-----GlyCysProGlyPro	: :	196
QY	267	CGATGTTTCATCAAGCCGATCTATCGGGGTGTCAGGCCCTGGCTCACCAAGTAGACCACT	: :	208
Db	197	-----ArgGlyGlu	: :	207
QY	207	CCTGCAGCAAGCTGGTGGCGGACTTGGGGTAGGTGACGATCCACACGTGCTGGGCGCA	: :	148
Db	208	GlnAla	: :	213
QY	147	CCGGGAAGTTGGGATCTCTCCATCTTCCCGGGCGCAGAAAGGGCGGCGCACGCCAT	: :	88
Db	214	ProProGlyProArgGlyProProGlyProGluGlyAsnProGlyGlyAlaGlyGluAsp	: :	233
QY	87	GGAACTCGAAGTACTGTCTCTCGAATCCCGGGGTGTGGGGTCTCGGCTCGCTCT	: :	28
Db	234	GlyAsnGlnGlyProValGlyHis	: :	250
QY	27	CCG	: :	25
Db	251	Pro	: :	251
RESULT 79				
T43481				
probable mucin DKFPz434C196.1 - human (fragment)				
N;Alternate names: protein DKFPz434B0635.1				
C;Species: Homo sapiens (man)				
C;Date: 21-Jan-2000 #sequence revision 21-Jan-2000 #text_change 09-Jul-2004				
C;Accession: T43481; T34549; T17264				
R;Koehrer, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.				
submitted to the Protein Sequence Database, December 1999				
A;Reference number: Z22514				
A;Accession: T43481				
A;Molecule type: mRNA				
A;Residues: 1-580 <AAA>				
A;Cross-references: UNIPROT:Q9UF83; UNIPARC:UPI000006D400; EMBL:AL133561; NID:g6599133;				
A;Experimental source: adult testis; clone DKFPz434C196				
R;Poustka, A.; Weilenreuther, R.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.				
submitted to the Protein Sequence Database, October 1999				
A;Reference number: Z21540				
A;Accession: T34549				
A;Molecule type: mRNA				
A;Residues: 262-580 <POU1>				
A;Cross-references: UNIPARC:UPI0000070F36; EMBL:AL117481; NID:g6102864; PIDN:CAB59245.2;				
A;Experimental source: adult testis; clone DKFPz434B0635				
R;Poustka, A.; Klein, M.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.				
submitted to the Protein Sequence Database, September 1999				
A;Reference number: Z18723				
A;Accession: T17264				
A;Molecule type: mRNA				
A;Residues: 262-580 <POU2>				
A;Cross-references: UNIPARC:UPI0000070F36; EMBL:AL117481; NID:g5911958; PIDN:CAB55954.1;				
A;Experimental source: adult testis; clone DKFPz434B061				
C;Genetics:				
A;Note: DKFPz434C196.1; DKFPz434B0635.1; DKFPz434B061.1				
Alignment Scores:				
Pred. No.:	0.00469	Length:	580	
Score:	148.50	Matches:	76	
Percent Similarity:	33.1%	Conservative:	17	
Best Local Similarity:	27.0%	Mismatches:	85	

RESULT 80

JQ0317

hypothetical 82K protein - Xanthomonas campestris pv. vesicatoria

C;Species: Xanthomonas campestris pv. vesicatoria

C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004

C;Accession: JQ0317

R;Bonas, U.; Stall, R.B.; Staackiewicz, B.

Mol. Gen. Genet. 218, 127-136, 1989

Query Match:	3.4%	Indels:	103	
DB:	2	Gaps:	14	
US-10-768-158-1 (1-2419) x T43481 (1-580)				
QY	789	AGATGTCCTTCCACAGCCCAACTCTTC	-----CCCGGCCACGGCAGGG-----	745
Db	71	ArgThrProThrArgAlaSerLeuMetArgThrProProArgAlaSerProThrArgLys	: :	90
QY	744	CCTCAGCGTTGACAGCACTGGTCCACAGCTGGTGGCAGT	-----GCTCCGCTCA	697
Db	91	ProProArgAlaSerProArgThrProSerArgAlaSerProThrArgArgLeuProArg	: :	110
QY	696	GGGCTTCAGCTGGGCTTGTAC	-----AGGACACCCAGGAATCTGG	652
Db	111	AlaSerProMetGlySerProHisArgAlaSerProMetArgThrProProArgAlaSer	: :	130
QY	651	CCAGCTGCTCCACCATCGTCACCAAGTCCGATGTCCTCATCTTCTGAGAAAAGCA	: :	592
Db	131	ProThrGlyThrProSerThrAlaSerPro	-----ThrGlyThrProSer	145
QY	591	CGTTCGAGTCCATGCGGTGCTCCC	-----AGAACTCCTGCACGTGCT	550
Db	146	SerAlaSerProThrGlyThrProProArgAlaSerProThrGlyThrProProArgAla	: :	165
QY	549	CAAAACAGGAGCGGTAGCCAGCTTATCATTTATAAACCCTCCGGCAGAAATCTTGAAGG	: :	490
Db	166	TrpAla--ThrArgSerProSerThrAlaSerLeuThr	: :	182
QY	489	TGCCTCGGTAGCTCATGGTCCGACAGAGCGGTGGAACTGATATAAGACACCAACAGAT	: :	430
Db	183	AlaSerLeuThrArgTrpProProArgAlaSerProThr	-----ArgThrProPro----	199
QY	429	CCTTGGGGTTGCGAGCCATATAGATGACCTTGGAGTCTCCATTGTGGAGGTCCAGAGGCA	: :	370
Db	199	-----	-----	199
QY	369	GAAGCGGTAGGCGAGGTGGCTTGTGATGAGCGGGGAGAGTTCAGTTCCTTGATGATGT	: :	310
Db	199	-----	-----	199
QY	309	CCAGGCCCGGTGTGGGTACTCCAGGACCGGGAGCTGCTCGTGCATGTTTCATCAAGCCGA	: :	250
Db	200	-----	-----ArgGluSerProArgMet	205
QY	249	TCTCATCGGGTTCAGCGCCCTGGCTCACCAAGTAGACCACTCCTCGCAGCAAGCTGTGC	: :	190
Db	206	SerHisArg	-----AlaSerProThrArgThrProProArgAlaSer	219
QY	189	CGGACTTGGGGTAGGTACGATCCACACGTGCTGGGCGCCACCCGGGAAAGTTGGCGATCT	: :	130
Db	220	-----	-----ProThrArgArgProProArgAlaSerProThrArgThr	232
QY	129	CCTCCATCTTCCCGCGGAGCA	-----AGGGCGGAGCGCCAGCCCATGGAACCTCGA	79
Db	233	ProProArgGluSerLeuArgThrSerHisArgAlaSerProThrArgMetProProArg	: :	252
QY	78	AGTACTTGCTCTCGAACTCCCGGGGTGC	-----TGGGGGTCTCGGCTCGCTCTCGG	25
Db	253	AlaSerProThrArgArgProProArgAlaSerProThrGlySerProProArgAlaSer	: :	272
QY	24	CCA	22	
Db	273	Pro	273	
RESULT 80				
JQ0317				
hypothetical 82K protein - Xanthomonas campestris pv. vesicatoria				
C;Species: Xanthomonas campestris pv. vesicatoria				
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004				
C;Accession: JQ0317				
R;Bonas, U.; Stall, R.E.; Staskawicz, B.				
Mol. Gen. Genet. 218, 127-136, 1989				

A:Title: Genetic and structural characterization of the avirulence gene avrBs3 from Xanth

A:Reference number: JQ0316; MUID:89384426; PMID:2550761

A:Accession: JQ0317

A:Molecule type: DNA

A:Residues: 1-784 <BON>

A:Cross-references: UNIPROT:PI4728; UNIPARC:UPI00013A178

C:Comment: X. campestris pv. vesicatoria is the causal agent of bacterial leaf spot dise

C:Genetics:

A:Start codon: GTG

Alignment Scores:

Pred. No.: 0.005 Length: 784
Score: 148.50 Matches: 152
Percent Similarity: 30.1% Conservative: 58
Best Local Similarity: 21.8% Mismatches: 238
Query Match: 3.4% Indels: 249
DB: 2 Gaps: 40

US-10-768-158-1 (1-2419) x JQ0317 (1-784)

```
QY 58 GGGAGTTCAGAGCAAGTACTTCGAGT-----TCCATGGCG 93
Db 137 GlyAsnSerArgCysThrValSerSerAlaCysLeuProSerTrpLeuAlaMetAla 156
QY 94 TCGGGCTGCGCCCTCTCGCGGGGAAGATGGAGGAGATCGCAACTTCCCGGTGCGGC 153
Db 157 ThrThrCysSerGlyValArgProTrpAlaTrpHisSerThrGlyAsnSerArgCys--- 175
QY 154 CCAGCGCATGTGGATCGTCACTACCCCAAGTCCG-----GCACGAGTTGC 201
Db 176 -----ThrValSerSerAlaGlyLeuProProLeuLeuAlaMetAlaThrThrCys 193
QY 202 TGC-----AGAGGTGGTCT-----ACTGGTGAGCC 228
Db 194 CysGlyValArgProTrpAlaTrpHisSerThrGlyAsnSerArgCysThrValSerSer 213
QY 229 AGGGCGCTGACCCGATGAGATCGGCTTGATGAACATCGACGAGCAGCTCCCGTCTCG 288
Db 214 AlaCysLeuProSerTrpLeuAlaMetAlaThrThrCysSerGlyValArgProTrp 233
QY 289 AGTACCACACGCGCGCTCGGATCATCAAGAACTGACCTCCCGCGCTCATCAAGA 348
Db 234 -----AlaTrpHisSerThrGlyAsnSerArgCysThrValSerSer--- 247
QY 349 GCACCTGCGCTTACCGCTTCTCCCTCTGACCTCCACATGAGAGACTCAAGGTCTATCT 408
Db 248 AlaCysLeuProSerTrpLeuAlaMetAlaThrThrCysSerGlyValArgProTrp 267
QY 409 ATATGGCTCGCAACCCCAAGGATCTGGTGTGCTTATTATCATGATTCCACCGCTCTGC 468
Db 268 AlaTrpHisSerThrGlyAsnSer---ArgCysThrValSerSerAla---Cys 283
QY 469 GGACCATGAGTACCGAGGCACCTTTCAAGAATTCTGCGGAGGTTTATGAATGATAAGC 528
Db 284 LeuPro-----ProSer 287
QY 529 TGGGCTAGCGCTCTCGTTTGACACATGTCGAGAGTTCTGGGAGCAGCGATCGACTCGA 588
Db 288 TrpLeuAlaMet-----AlaThrThrCysSerGlyValArgProTrpAlaTrpHisSer 305
QY 589 ACG-----TGCCTTTTCTCAAGT----- 606
Db 306 ThrGlyAsnSerArgCysThrValSerSerAlaCysLeuProProLeuLeuAlaMet 325
QY 607 ATGAAGCATGTCATCGGACCTGGTACCATGGTGGAGCAGCTGGCGAGATTCTCGGGGG 666
Db 326 AlaThrThrCysSerGlyValArgProTrpAlaTrpHisSerThrGlyAsnSer---Ala 344
QY 667 TGTCTGTGACAAAGGCCACCTGGAAGCCCTGAGGAGCACTGCCACCACTGGTGGACC 726
Db 345 CysThrValSerSerAlaCysLeuProPro-----LeuLeuLeuAlaMetAlaThr 361
QY 727 AGTGCTGCAACGCTGAGGCCCTCGCCGTGGCGGGGAAGAGTGGGCTGTGGAAGGACA 786
```

```
Db 362 ThrCysSerGlyValArg-----ProTrpAlaTrpHisSerThrGlyAsnSerArgCys 379
QY 787 TCT-----TCACCGTCTCATGAATGAGAGTTTGACTTGGTGATATAAC 831
Db 380 ThrValSerSerAlaCysLeuProProLeuLeuAlaMetAlaThrThrCysCysGly 399
QY 832 AGAAGA-----TGGGAAGTGTGACCTCACGTTTGACTTTTATTTATATAACACA 882
Db 400 ValArgProTrpAlaTrpHisSerThrGly-----AsnSerArg 412
QY 883 AACAAACAACCTGCTCATACATACCCAGACAGTCTACTAGCCAAAAGTCTCTGATGCA 942
Db 413 -CysThrValSerSerAlaCysLeu-ProProSerTrpLeuAlaMetAlaThrThrCys- 431
QY 943 TTCATTATTCTCTGCTGACAAACTCTGGAAGCAGCGTGTGAAACAGCGGGGGAAGGA 1002
Db 431 ----- 431
QY 1003 AGAGCGCGTGGAGCGGAGTGTGATGATTCCCAACCGAAGCAGCTGTCTCGCCTTTA 1062
Db 432 --SerGlyValArgProTrpAlaTrpHisSerThrGlyAsnSerArgCysThrValSerS 451
QY 1063 GAACGTGACGCTCTCCATGCTCTGATTACAAACAGTCTCCACATTGCACT----- 1112
Db 451 erAlaCys--LeuProProSerTrpLeuAlaMetAlaThrThrCysSerGlyValArgP 470
QY 1113 -----TCCATGGCTGGACCGTAAGATAAAGCTGTATATATATGCAACT 1158
Db 470 roTrpAlaTrpHisSerThrGly-----AsnSerAlaCysThrValSerSer- 485
QY 1159 AGAATGCTGCTGCTTTTCAACCCCGTATTATTGTTATTT-----TATAG 1200
Db 486 -AlaCysLeuPro-----ProLeuLeuAlaMetAlaThrThrCysSerGlyValAr 503
QY 1201 AGCTTTTCACTGAAATCTACATAAATGTACATAAACCAAAATAAAGATTTCATTTCOA 1260
Db 503 gProTrpAlaTrpHisSerThr-----G 511
QY 1261 GGAATCAGGAGGAGCCACACCGGAATGGTAGAAGATCTCAGGGTTAACTCTTTATTT 1320
Db 511 yAsnSerAlaCysThrValSerSerAlaCys-----LeuPr 523
QY 1321 TGTAGTTTATTATCTAAGGCACAGCATTTCTTCTCACTTGGTTCTCAGATAGTGGTG 1380
Db 523 oProLeuLeuAlaMetAlaThrThr---CysSerGly-----ValArgProTrp-- 539
QY 1381 AGAACAGAGGATGAGTTGGTCTGTTTGGGGGAATCTGGACACTTGTTTTATTCTGAC 1440
Db 540 -----AlaTrpHisSerThrGly----- 545
QY 1441 GTTCACTTCTCAGAACCTTCTCGAATGAGCAGAAATTTGTTCACTAGTCTTCAGATG 1500
Db 546 -----AsnSerAlaCy 549
QY 1501 GAGCTCTTCTGCCAGAGACTTCCAGCGCGCTCCAAAGGCCCAATGCAGAGGAGCCC 1560
Db 549 sThrValSerSerAlaCysLeuPro-----ProLeuLeuAlaMe 563
QY 1561 GCGGACATGTGCTGAGGGAAGTCTGCTCTGAGGCTGGCAGGTGGGAGTCTAATGCAG 1620
Db 563 tAlaThrThrCysSerGlyValArgProTrp--AlaTrp----- 575
QY 1621 TCAGGAGCATTTGCATGTCAGTGGTGGAGAGTGGCGCCACCAAGGACCGAGTTGCGCTC 1680
Db 576 -----HisSerThrGlyAsnSerArgCysThrValSerSerAlaCys----- 589
QY 1681 GAATTTGAGTGAATTCACAGCTTACTTTGTTCTTGAAGTGATAGCTACTAATGCT 1740
Db 590 -----LeuProProPro-----LeuLeuAlaMe 597
QY 1741 GGCA-----AGCAGATG 1752
```

Db	597	talaThrThrCysSerGlyValArgProTrpAlaTrpHisSerThrGlyAsnSerArgCy	617
QY	1753	CTTAATAGTAATTTCTAAATCCCGGGTCTTTATCATTCAGTTGTTCTGTGCACCTG	1812
Db	617	sthrValSerSerAlaCysLeuProProLeuLeuAlaMetAlaThrThrCysCys--	636
QY	1813	AGCGCGTCAGCGTGGAGAGACCATTTTGGAGTGTAGCCCTGTTTCATCTCGATCAGGT	1872
Db	637	-GlyValArgProTrpAlaTrpHis-----SerTh	646
QY	1873	TGCGACGGCGGCTGC--GTGCTGTCCACCTCATCCCTCCG	1912
Db	646	rGlyAsnSerArgCysThrValSerSerAlaCysLeuProPro	660
RESULT 81			
A40020			
collagen alpha 1(XII) chain precursor - chicken			
N:Alternate names: fibrochimerin			
C:Species: Gallus gallus (Chicken)			
C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text change 09-Jul-2004			
C:Accession: A40020; A34485; A28037; S23814; S23254; S28811			
R:Yamagata, M.; Yamada, K.M.; Yamada, S.S.; Shinomura, T.; Tanaka, H.; Nishida, Y.; Obay			
J. Cell Biol. 115, 209-221, 1991			
A:Title: The complete primary structure of type XII collagen shows a chimeric molecule w			
nous region of type IX collagen, and short collagenous domains with an Arg-Gly-Asp site.			
A:Reference number: A40020; MUID:92011862; PMID:1918137			
A:Accession: A40020			
A:Molecule type: mRNA			
A:Residues: 1-3124 <YAM>			
A:CROSS-references: UNIPROT:P13944; UNIPARC:UPI0000136D2D; GB:D00824; NID:G222810; PIDN:			
A:Note: In the authors' translation residues 1216-1219 are shown after residue 1235 and,			
R:Gordon, M.K.; Gerecke, D.R.; Dublet, B.; van der Rest, M.; Olsen, B.R.			
J. Biol. Chem. 264, 19772-19778, 1989			
A:Title: Type XII collagen. A large multidomain molecule with partial homology to type I			
A:Reference number: A34485; MUID:90622079; PMID:2584192			
A:Accession: A34485			
A:Molecule type: mRNA			
A:Residues: 2456-2758, 'A', 2760-2802, 'F', 2804-2976, 'F', 2978-3124 <GOR>			
A:CROSS-references: UNIPARC:UPI0000171233; EMBL:J051137; NID:G211284; PIDN:AAA48635.1; PI			
A:Accession: B34485			
A:Molecule type: protein			
A:Residues: 2772-2792;2846-2873 <GOR2>			
A:CROSS-references: UNIPARC:UPI0000173C3F; UNIPARC:UPI0000173C40			
R:Gordon, M.K.; Gerecke, D.R.; Olsen, B.R.			
Proc. Natl. Acad. Sci. U.S.A. 84, 6040-6044, 1987			
A:Title: Type XII collagen: distinct extracellular matrix component discovered by cDNA c			
A:Reference number: A28037; MUID:87317590; PMID:3476925			
A:Accession: A28037			
A:Molecule type: mRNA			
A:Residues: 2960-2976, 'F', 2978-3074, 'AG' <GOR3>			
A:CROSS-references: UNIPARC:UPI00001712P4; EMBL:M17375; NID:G211649; PIDN:AAA48718.1; PI			
A:Note: this sequence has been revised in reference A34485			
R:Koch, M.; Bernasconi, C.; Chiquet, M.			
Eur. J. Biochem. 207, 847-856, 1992			
A:Title: A major oligomeric fibroblast proteoglycan identified as a novel large form of			
A:Reference number: S23814; MUID:92362621; PMID:1323460			
A:Accession: S23814			
A:Molecule type: protein			
A:Residues: 'X',1333,'Q',1335-1347;1914-1928;2504,'X',2506,'X',2508-2511,'X',2513-2517 <			
A:CROSS-references: UNIPARC:UPI0000173C41; UNIPARC:UPI0000173C42; UNIPARC:UPI0000173C43			
R:Dublet, B.; van der Rest, M.			
J. Biol. Chem. 262, 17724-17727, 1987			
A:Title: Type XII collagen is expressed in embryonic chick tendons. Isolation of pepsin-			
A:Reference number: S22254; MUID:88087065; PMID:3121603			
A:Accession: S22254			
A:Molecule type: protein			
A:Residues: 2831-2832,'T',2834,'R',2836-2843;3002-3014 <DUB>			
A:CROSS-references: UNIPARC:UPI0000173C44; UNIPARC:UPI0000173C45			
R:Trubeb, J.; Trueb, B.			
Biochim. Biophys. Acta 1171, 97-98, 1992			
A:Title: The two splice variants of collagen XII share a common 5' end.			
A:Reference number: S28811; MUID:93042014; PMID:1420368			
A:Accession: S28811			

A>Status: preliminary			
A:Molecule type: mRNA			
A:Residues: 1-24,1189-1257,'S',1259-1263,'E',1265-1280 <TRU>			
A:CROSS-references: UNIPARC:UPI0000173C46; EMBL:X67327			
C:Genetics:			
A:Introns: 2845/3; 2863/3; 2887/3; 2899/3; 2922/1; 2985/1; 3008/1; 3065/1			
C:Keywords: alternative splicing; cell binding; coiled coil; connective tissue; disulfid			
F:1-23/Domain: signal sequence #status predicted <SIG>			
F:24-3124/Product: collagen alpha 1(XII) chain #status predicted <MAT>			
F:24.1189-3124/Product: collagen alpha 1(XII) chain short splice form #status predicted			
F:24-114/Domain: IIIA #status predicted <IIIA>			
F:24-105/Domain: fibronectin type III repeat homology <FN3A>			
F:137-301/Domain: von Willebrand factor type A repeat homology <VWA1>			
F:332-425/Domain: IIIB #status predicted <IIIB>			
F:437-601/Domain: fibronectin type III repeat homology <FN3B>			
F:629-1178/Domain: von Willebrand factor type A repeat homology <VWA2>			
F:630-711/Domain: IIIC #status predicted <IIIC>			
F:721-802/Domain: fibronectin type III repeat homology <FN3C>			
F:812-895/Domain: fibronectin type III repeat homology <FN3D>			
F:905-986/Domain: fibronectin type III repeat homology <FN3E>			
F:995-1076/Domain: fibronectin type III repeat homology <FN3F>			
F:1086-1169/Domain: fibronectin type III repeat homology <FN3G>			
F:1197-1361/Domain: fibronectin type III repeat homology <FN3H>			
F:1384-2295/Domain: von Willebrand factor type A repeat homology <VWA3>			
F:1384-1465/Domain: IIID #status predicted <IIID>			
F:1474-1557/Domain: fibronectin type III repeat homology <FN3I>			
F:1566-1647/Domain: fibronectin type III repeat homology <FN3K>			
F:1655-1738/Domain: fibronectin type III repeat homology <FN3L>			
F:1756-1838/Domain: fibronectin type III repeat homology <FN3M>			
F:1847-1928/Domain: fibronectin type III repeat homology <FN3N>			
F:1937-2019/Domain: fibronectin type III repeat homology <FN3O>			
F:2028-2110/Domain: fibronectin type III repeat homology <FN3P>			
F:2119-2199/Domain: fibronectin type III repeat homology <FN3Q>			
F:2207-2294/Domain: fibronectin type III repeat homology <FN3R>			
F:2325-2490/Domain: von Willebrand factor type A repeat homology <VWA4>			
F:2438-2440/Region: cell adhesion #status predicted			
F:2509-2750/Domain: IXP, homologous to NC4 domain of type IX collagen #status predicted			
F:2751-2902/Domain: collagenous COL2 #status predicted <COL2>			
F:2895-2901/Region: cell attachment (R-G-D) motif			
F:2903-2945/Domain: non-collagenous NC2 #status predicted <NC2>			
F:2946-3048/Domain: collagenous COL1 #status predicted <COL1>			
F:3049-3124/Domain: non-collagenous NC1 #status predicted <NC1>			
F:32.1006,1032,1044,1512,1767,2210,2273,2532,2683/Binding site: carbohydrate (Asn) (cova			
F:2780,2789,2836,2842,2860,2866,2869,3004,3007/Modified site: hydroxyproline (Pro) #stat			
Alignment Scores:			
Pred. No.:	0.00738	Length:	3124
Score:	148.00	Matches:	87
Percent Similarity:	31.3%	Conservative:	13
Best Local Similarity:	27.3%	Mismatches:	107
Query Match:	3.4%	Indels:	112
DB:	2	Gaps:	21
US-10-768-158-1 (1-2419) x A40020 (1-3124)			
QY	762	CCCGGCCACCGGCGGCGCTCAGGTTGTCAGCACTGTTCCACCAGCTGGTGGCAGTGCT	703
Db	2756	ProGlyProProGlyGlyProGlyAlaLysGlyProArgGlyGluArgGlyLeuThrGly	2775
QY	702	CGTCAGGCGCTTCCAGCTGGGCGCTTGTCCACAGGACA---CCCCAGGAATCTGGCCAGCT	646
Db	2776	SerSerGly---ProProGlyProArgGlyGluThrGlyProProGly-----ProGln	2792
QY	645	GCTCCACCATCGTCCACAGGTCGCCGATGCATGCTCTTCTACTTGTAGAAAAACACAGT---	589
Db	2793	GlyProProGlyProGlnGlyProAsnGlyLeuGlnIleProGlyGluProGlyArgGln	2812
QY	588	-----TCGAGTCCATCGGTCCTCCAGAACTCTCGCACGTCGTCACACGAGGAGC---	538
Db	2813	GlyMetLysGlyAspAlaGlyGlnProGlyLeuProGlyArgSerGlyThrProGlyLeu	2832
QY	537	-----CGTAGCCCACTTATCATTAATAAACCTCCGCGAGAAATCTTTGAAAGGTGC	487


```

Db      2833 ProGlyProProGlyProVal-----GlyProProGlyGluArg--- 2845
QY      486 CTCGGTAGCTCATGTCCGACAGAGCGGTGAATGATAATAAGACACACAGATCCT 427
Db      2846 -----GlyPheThrGlyLysAspGlyProProGlyProProGlyPro 2862
QY      426 TGGGTTCCGAGCCATATAGATGACCTTGGAGTCTCCATTGGAGGTGACAGGGCAGAA 367
Db      2863 AlaGlyAlaPro-----GlyValProGlyValAlaGlyPro 2874
QY      366 AGCGGTAGGCGAGGTGGCTCTTTGATGAGCGGGAGAGGTGAGTCTCTTGTGATGTCCA 307
Db      2875 SerGlyLysProGlyLysPro-----GlyAspArgGlyThrPro----- 2887
QY      306 GCCCGCGTGGGTGACTCCAGGA----- 283
Db      2888 -----GlyThrProGlyMetLysGlyGluLysGlyAspArgGlyAspIleAla 2903
QY      282 -----CCGGAGCTGCTGCTCGA---TGTTTCATCA----- 256
Db      2904 SerGlnAsnMetMetArgAlaValAlaArgGlnValCysGluGlnLeuIleAsnGlyGln 2923
QY      255 ---AGCCGA-----TCTCATCGGGGT 238
Db      2924 MetSerArgPheAsnGlnMetLeuAsnGlnIleProAsnAspTyrTyrSerAsnArgAsn 2943
QY      237 CAGCCGCTGCTGCTACCAAGTAGACCA---CCTCTGCGACGAGTGTGTGCGGACTTGG 181
Db      2944 Gln---ProGlyProProGlyProProGlyProProGlyAlaAla----- 2957
QY      180 GTAGGTGACGATCCACAGTCCGTCGGCGCCGACCGGA----- 142
Db      2958 -----GlyThrArgGlyGluProGlyProGlyProGlyArgProGlyPheProGly 2973
QY      141 -----AGTTGGCGATCTCTCCATCTTCCCGCGAGAGGCGCGACCGCCGCAT 88
Db      2974 ProProGlyValGlnGlyProProGlyProGlyGluArgGlyMetProGlyGluLysGlyGluArg 2993
QY      87 GGAATCGAAGTACTTGC-----TCT 67
Db      2994 GlyThrGlySerGlnGlyProArgGlyLeuProGlyProProGlyProGlnGlyGluSer 3013
QY      66 CGAATC---CCCCGGGTGCTGGGGTCTCGG---CCTCGCTCTCCCGCATGCGCGC 16
Db      3014 ArgThrGlyProProGlySerThrGlySerArgGlyProProGlyProProGlyArg 3032

RESULT 82
T13049
e:elid - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C>Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C:Accession: T13049
R:Trelisman, J.E.; Luk, A.; Rubin, G.M.; Heberlein, U.
submitted to the EMBL Data Library, March 1998
A:Reference number: Z17592
A:Accession: T13049
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-2715 <TRE>
A:Cross-references: UNIPROT:Q8IN94; UNIPARC:UPI000007D87E; EMBL:AF053091; NID:g2981220;
C:Genetics:
A:Gene: eld
A:Cross-references: FlyBase:FBgn0003013
C:Function:
A:Description: could act as a transcription factor antagonistic to the Wg pathway
C:Keywords: DNA binding

Alignment Scores:
Pred. No.: 0.00782 Length: 2715
Score: 147.50 Matches: 82
Percent Similarity: 36.9% Conservative: 36
Best Local Similarity: 25.6% Mismatches: 98

```

```

Query Match: 3.4% Indels: 104
DB: 2 Gaps: 16

US-10-768-158-1 (1-2419) x T13049 (1-2715)
QY 759 GGCACAGCGGAGGGCTCAGCGTTGCAGCACTGCTCCACCACTGTTGGCAGTGTCT--- 703
Db 1433 GlyProGlnProProThrGlnThrValAlaGlyProProAlaGlyGlyAlaProGly 1452
QY 702 -----CCGTCAGGGCTTCAGCTGGGCTTGTGCACAGACACCC 664
Db 1453 AlaProProSerSerAlaTyrProThrGlyArgProSerGlnGlnAspTyrTyrGlnPro 1472
QY 663 CCA----- 661
Db 1473 ProProAspGlnSerProGlnProArgArgHisProAspPheIleLysAspSerGlnPro 1492
QY 660 -----GGAATCTGCCAGCTGTCTCCA----- 640
Db 1493 TyrProGlyTyrAsnAlaArgProGlnIleTyrGlyAlaTyrGlnSerGlyThrGlnGln 1512
QY 639 -----CATGCTCACCAGGTCCCGATGCATGCTCTTCATCTTGAGAAAAA 595
Db 1513 TyrArgProGlnTyrProSerSerProAlaProGlnAsnTrpGlyGlyAlaProProArg 1532
QY 594 GCACGTTCCAGTCCATCGGTGCTCCAGAACTCTCGACGCTGCTCAAAACAGGAGCCGT 535
Db 1533 GlyAlaAlaProProProGlyAlaProHisGlyProPro-IleGlnGlnProAlaGlyVa 1552
QY 534 AGCCGAGCTTATCATTAACCTCCGGCAGAAATTTCTGAAAGGTGCTCGTAGCTCA 475
Db 1552 IAlaGlnTrpAspGlnHisArgTyrProProGln-----G1 1564
QY 474 TGGTCCG-----CAGAGAGCGTGGAACTGATGATAATAAGACCA 436
Db 1564 nGlyProProProProGlnGlnGlnGlnProGlnGlnGlnGlnGlnGlnGlnGlnPro 1583
QY 435 CCA-----GATCCTTGGGTTGCGAGCCATATAGTACCT---TGG 397
Db 1584 ProTyrGlnGlnValAlaGlyProProGlyGlnGlnProGlnAlaProProGlnTrp 1603
QY 396 AGTCTCCATTGTGAGGTGAGAGGCGAGAAAGCGTAGGCGAGGTGCTCTTGATAGGC 337
Db 1604 AlaGlnMetAsnProGlyGlnThrAlaGlnSerGlyIleAlaPro----- 1618
QY 336 GGGGAGAGGTGAGTCTTGTGATGATGTCAGCGCCGCTGTTGGTACTCCA---GGACGG 280
Db 1619 -----ProGlySerProLeuArgProProSerGlyPro 1629
QY 279 GGAGCTGCTCGTATGTTTCATCA-----AGCCGATCTCATCGGGGTGAGCGCCCTGC 226
Db 1630 GlyGlnGlnAsnArgMetProGlyMetProAlaGlnGlnGlnGlnSerGlnGlnGlnGly 1649
QY 225 TCACCAAGTAGACACACCTCTCCAGCAAGCTGGTCCGAGCTTGGGTAGGTGACGATCC 166
Db 1650 GlyValProGlnProProProGlnGlnAlaSerHisGly---GlyVal-----Pr 1665
QY 165 ACAGTCTCGTGGCGGCGACCGGGAAGTTGGGATCTCTCCATCTTCCTCGCGGCGAGAAG 106
Db 1665 oSerProGlyLeuPro-----GlnValGly-----ProGlyGlyMetVa 1678
QY 105 GCGGAGCGCGCACGCCATG-----GAACTCGAAGTACTTGTCTCTCGAACTCCCCG 55
Db 1678 lLysProProTyrAlaMetProProProProSerGlnGlyValGlyGlnGlnValGlyGl 1698
QY 54 GGGTGTCTGGGGTCTCGGCTCTCGCTCTCCGCCATCGCGCCGCGCGCGCGCG 5
Db 1698 nGlyProProGly---GlyMet-MetSerGlnLysProProMetPro 1713

RESULT 83
PIHUSD
salivary proline-rich glycoprotein precursor PRB4 (large allele) [validated] - human
N:Contains: basic proline-rich protein IB-5; proline-rich peptide P-D

```


C:Species: Homo sapiens (man)
 C:Date: 19-Feb-1984 #sequence_revision 12-Apr-1996 #text change 05-Oct-2004
 C:Accession: S03176; S03175; S10890; D25372; E38355; A03295; A61294; S62891
 R:Lyons, K.M.; Stein, J.H.; Smithies, O.
 Genetics 120, 267-278, 1988
 A:Title: Length polymorphisms in human proline-rich protein genes generated by intragenic recombination
 A:Reference number: S02127; MUID:89121440; PMID:2851479
 A:Accession: S03176
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Cross-references: UNIPARC:UPI0000174519; UNIPARC:UPI0000174516; EMBL:X07704
 A:Note: large allele
 A:Accession: S03175
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 35-310 <LY1>
 A:Cross-references: UNIPROT:P10163; UNIPROT:P10162; UNIPARC:UPI0000132351; EMBL:X07715
 A:Note: medium allele
 A:Accession: S10890
 A:Status: preliminary; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-38, 60-112, 'T', 114-115, 'P', 117-121, 185-271, 'A', 273-310 <LY3>
 A:Cross-references: UNIPARC:UPI000016A5FF; EMBL:X07882; NID:935647; PIDN:CAA30729.1; PID:R:Maeda, N.; Kim, H.S.; Azen, E.A.; Smithies, O.
 J. Biol. Chem. 260, 11123-11130, 1985
 A:Title: Differential RNA splicing and post-translational cleavages in the human salivary proline-rich protein
 A:Reference number: A92492; MUID:85289325; PMID:2993301
 A:Accession: D25372
 A:Molecule type: mRNA
 A:Residues: 1-36, 'E', 38-112, 'T', 114-115, 'P', 117-121, 185-271, 'A', 273-310 <MAE>
 A:Cross-references: UNIPARC:UPI000013DBDC
 R:Kaufman, D.L.; Bennick, A.; Blum, M.; Keller, P.J.
 Biochemistry 30, 3351-3356, 1991
 A:Title: Basic proline-rich proteins from human parotid saliva: relationships of the covariation of amino acid sequence with molecular weight
 A:Reference number: A38355; MUID:91190884; PMID:1849422
 A:Accession: E38355
 A:Molecule type: protein
 A:Residues: 241-254, 'KN', 257-310 <KAU>
 A:Cross-references: UNIPARC:UPI0000174517
 R:Saitoh, E.; Iseura, S.; Sanada, K.
 J. Biochem. 93, 495-502, 1983
 A:Title: Complete amino acid sequence of a basic proline-rich peptide, P-D, from human parotid saliva
 A:Reference number: A03295; MUID:83186122; PMID:6841349
 A:Accession: A03295
 A:Molecule type: protein
 A:Residues: 241-310 <SAI>
 A:Cross-references: UNIPARC:UPI0000174518
 R:Shimomura, H.; Kanai, Y.; Sanada, K.
 J. Biochem. 93, 857-863, 1983
 A:Title: Amino acid sequences of glycopeptides obtained from basic proline-rich glycoproteins
 A:Reference number: A61294; MUID:83265671; PMID:6874667
 A:Accession: A61294
 A:Molecule type: protein
 A:Residues: 54-57, 'E', 59-73, 'R', 82-101 <SHI>
 A:Cross-references: UNIPARC:UPI0000174519; UNIPARC:UPI000017451A
 R:Charlton, A.J.; Baxter, N.J.; Lilley, T.H.; Haslam, E.; McDonald, C.J.; Williamson, M.
 FEBS Lett. 382, 289-292, 1996
 A:Title: Tannin interactions with a full-length human salivary proline-rich protein disulfide isomerase
 A:Reference number: S62891; MUID:96184506; PMID:8605987
 A:Accession: S62891
 A:Molecule type: protein
 A:Residues: 241-252 <CHA>
 A:Cross-references: UNIPARC:UPI000017451B
 A:Note: amino end of peptide designated basic proline-rich protein IB-5
 A:Note: it is unclear from the peptide sequence whether this is a product of the PRB2 (P) or PRB1 (P)
 C:Genetics:
 A:Gene: GDB:PRB4
 A:Cross-references: GDB:119514; OMIM:180990
 A:Map position: 12p13.2-12p13.2
 A:Introns: 22/1; 34/1
 A:Note: the list of introns may be incomplete
 C:Keywords: glycoprotein; saliva; tandem repeat

F:1-16/Domain: signal sequence #status predicted <SIG>
 F:241-310/Product: proline-rich peptide P-D #status experimental <MAT>
 F:66,87,171/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F:108,150,192,213,234/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:
 Pred. No.: 0.00531 Length: 310
 Score: 147.00 Matches: 71
 Percent Similarity: 34.0% Conservative: 20
 Best Local Similarity: 26.5% Mismatches: 85
 Query Match: 3.3% Indels: 92
 DB: 1 Gaps: 17

US-10-768-158-1 (1-2419) x PIHUSD (1-310)

QY 762 CCGCGCCACCGGCGGCGCTTCAGCTTCACGACTGCTCCACGAGCTGGTGGCAGTGTCT 703
 DB 54 ProGlyLysProGlnGlyPro-----ProProGlnGlyGlyAsnGln 67
 QY 702 CCGTCAGGGCTTCAGCTGGGCGCTTGTACAGGACACCCAGGAATCTGGCCAGCTGCT 643
 DB 68 SerGlnGlyProPro-----ProProProGlyLysProGluGlyArg 81
 QY 642 CCACCA-----TCGTCCACAGGTCCCGATGCTCTTCATACCTTGAGAAAAA 595
 DB 82 ProProGlnGlyGlyAsnGlnSerGlnGlyPro----- 93
 QY 594 GCACGTTCCAGTCCAGTGGCGTCCAGAACTCTCCGACGTGCTCAAAACAGGAGCGCT 535
 DB 94 -----ProHisProGlyLysProGluArgProProProGlnGlyGlyAsnGlnSerGln 111
 QY 534 AGCCAGCTTATCATATAAACCTCCGCGAGAAATCTTGAAGTGCTCGGTAGCTCA 475
 DB 112 GlyProProHis-----ProGlyLysProGluSerArg----- 123
 QY 474 TGGTCCGAGAGAGCGGTGGAAGTATGATGCTGGAGTCTCCATTTGGAGGTACAGAGGCGAGAA 367
 DB 142 GluGlyProProProGln-----GlyGlyAsnGlnSerGln 153
 QY 366 -----AGCGGTAGGCGAGGTGGCTCTTGATGAGCGGG 334
 DB 154 GlyThrProProProGlyLysProGluGlyArgProProGln-----GlyGly 170
 QY 333 GAGAGTTCAGTCTTGTGATGATGCTCCAGCGCGGTGTTGGTACTCCAGGACGGGAGCT 274
 DB 171 AsnGlnSer-----GlnGlyPro-----ProHisProGlyLys 182
 QY 273 GCTCGTCA-----TGTTTCATCAAGCGCATCTCATCGG---GGTCAGCGCCTGGC 226
 DB 183 ProGluArgProProProGlnGlyGlyAsnGlnSerHisArgProProProProGly 202
 QY 225 TCACCAAGTAGACCACTCTTCGACGAAGTGGTCCGAGCTGGGGTAGGTGAGCATCC 166
 DB 203 LysProGluArgProProProGln-----GlyGlyAsnGlnSer 215
 QY 165 ACACGTCTCGGCGCGCACCGGAAGTTGGCGATCTCTCCATCTTCCCGCGGAGAGG 106
 DB 216 GlnGlyProProHisProGlyLysProGluGlyProProProGlnGlnGlyAsnLys 235
 QY 105 GCGGAGCGCGCACCGCATGGAAGTCTGCTCTCAACTCCCGCGGGGTGCTGG 46
 DB 236 SerArgSerAla-----ArgSerProProGly----- 244
 QY 45 GGGTCTCGGCTGCTCTCCGCCA 22
 DB 245 -----LysProGlnGlyProPro 250

RESULT 84
 T24769

hypothetical protein T10B10.1 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T24769
R:Sim, M.
submitted to the EMBL Data Library, May 1996
A:Reference number: Z19934
A:Accession: T24769
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-428 <WIL>
A:CROSS-references: UNIPROT:Q22369; UNIPARC:UPI000007714C; EMBL:Z72514; PIDN:CAA96674.1;
A:Experimental source: clone T10B10
C:Genetics:
A:Gene: CESP:T10B10.1
A:Map position: X
A:Introns: 268/1

Alignment Scores:
Pred. No.: 0.0057 Length: 428
Score: 147.00 Matches: 81
Percent Similarity: 32.9% Conservative: 17
Best Local Similarity: 27.2% Mismatches: 112
Query Match: 3.3% Indels: 88
DB: 2 Gaps: 19

US-10-768-158-1 (1-2419) x T24769 (1-428)

```
QY 786 TGTCTCTCCACAGCCCACTTCCCGCGCCACCGGAGGCGCT---CAGCGTTGCAGC 730
      |||||::: ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 93 CysProAlaGlyAla-----ProGlyProProGlyAsnProGlyLysArgGlyAsp 109
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 729 ACTGTCACACAGCT-----CGTGGCAGTGTCCGTGAGGCTTCCAGCTGGGCTGTG 676
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 110 GluGlyHisProGlyAspGluGlyArgGlyAlaSerGlyIleSerLeuAlaThrThr 129
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 675 CACAGGACACCCCGAGGAATC-----TGGCCAGCTGTCTCCACCATCGTCACAGGTCCC 622
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 130 HisAspIleProGlyGlyCysIleLysCysProGluGlyProAlaGlyProProGlyPro 149
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 621 GATGATGTCTTCATACTTGAAGAAAGACCGTTCGAGTCCATCGGTGCTCCCGAGAACT 562
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 150 Asp-----GlyAspSerGlyProGluGlyPheProGlyLeu 161
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 561 CTGCAGCTGCTCAACACGAGGAGCGTAGCCAGCTTATCATATAACCTCCCGCAGA 502
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 162 GlnGlyGlnSer-----GlyProSerGlyGluAspGlyAlaProGlyGln 176
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 501 ATTCTTGAAGGTGCTCGTAGCTCATGTCGCGAGAGCGGTGGAACTCATATAAG 442
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 177 -----GluGlyAlaProGlyAspGlnGlyGluGlnGlyProLysGlyTyRAspGlyThr 194
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 441 ACACACCACAGAT----- 430
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 195 AspGlyProAspGlyGlnProGlyThrThrThrPheProGlyGlnAlaGlyGlnProGly 214
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 429 -----CCTTGGGTTCGAGCCATATAGATCACTTGAGTCTCATTTGGAGGTACG 376
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 215 GluProGlyTrpLeuGlyGluPro-----GlyLeuProGlyGln 227
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 375 AGGCAGCAAGCGGTAGGCGAGTGGCTCTGTATGAGCGCGGAGAGGTTCAGTTCCTTGA 316
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 228 HisGlyGluProGlyLysAspGlyGluGlu-----GlyProGlnGlyAlaPro--- 243
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 315 TGATGTCCAGGCCCGCT-----GTGGGTACTCCAGGA----- 283
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 244 GlyThrProGlyAsnAlaGlyHisAspAlaPheProGlyThrProGlyGlnAlaGlyLys 263
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 282 CCGGAGCT-----CCTGTCGATGTTATCAACCGCATCTCATCGGGTCAG 235
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 264 ProGlyAlaProGlyLysAspAlaAsnTyRProCysProGlnArgGlnAspAspArg 283
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 234 CGCCCTGGCTCACCAAGT-----AGACCACCTCTCTGCACAGAGCTGGTGCCGG 187
```

```
Db 284 ThrProProSerSerGlyThrSerAlaProGlnProProArgGlySer----- 300
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 186 ACTTGGGTAGGTACGATCCACAGTCGTCGCGCGCCGACCGGAGTTGGGATCTCT 127
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 -----ThrAlaAlaProGlyThr---ArgAlaPro 309
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 126 CCATCTTCCCGCGCGAGGCGGCGAGCCCGACCCCATCGAAGTACTTGCTCT 67
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 310 ProAlaThrArgAlaProProAlaThrArgAlaProProAlaThr-----Thr 325
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 66 CGAACTCCCCCGGGTGTCTGGCGGTCTCGCGCTCGCTCT---CGCCCATGCGCG 16
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 326 ArgAlaProProAlaThrThrArgProAlaProAlaSerGlnProProValArg 343
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

RESULT 85

S23810
collagen alpha 1(XVI) chain precursor - human
N:Alternate names: procollagen alpha 1(XVI) chain
C:Species: Homo sapiens (man)
C>Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jul-2004
C:Accession: S23810; PQ0612; S08012
R:Pan, T.C.; Zhang, R.Z.; Mattei, M.G.; Timpl, R.; Chu, M.L.
Proc. Natl. Acad. Sci. U.S.A. 89, 6565-6569, 1992
A:Title: Cloning and chromosomal location of human alpha1(XVI) collagen.
A:Reference number: S23810; MUID:92335339; PMID:1631157
A:Accession: S23810
A:Molecule type: mRNA
A:Residues: 1-1603 <PAN>
A:CROSS-references: UNIPROT:Q07092; UNIPARC:UPI0000126D33; EMBL:M92642; NID:g180757; PID
R:Yanaguchi, N.; Kimura, S.; McBride, O.W.; Hori, H.; Yamada, Y.; Kanamori, T.; Yamakoshi
J. Biochem. 112, 856-863, 1992
A:Title: Molecular cloning and partial characterization of a novel collagen chain, alpha
A:Reference number: PQ0612; MUID:93203161; PMID:1284248
A:Accession: PQ0612
A:Molecule type: mRNA
A:Residues: 'GGR', 421-536, 'P', 538-1159, 'P', 1161-1162, 'P', 1164, 'P', 1166-1603 <YAM>
A:CROSS-references: UNIPARC:UPI000016B3C2; GB:S57132; NID:g298641; PIDN:AAB25797.1; PID:
A:Experimental source: placenta
R:Kimura, S.
submitted to the EMBL Data Library, April 1989
A:Description: Partial nucleotide and amino acid sequence of a collagen-like protein fr
A:Reference number: S08012
A:Accession: S08012
A:Molecule type: mRNA
A:Residues: 403-419, 'GR', 421-536, 'P', 538-846, 'VM', <KIM>
A:CROSS-references: UNIPARC:UPI0000073DAB; EMBL:X14963; NID:g29984; PIDN:CAA33085.1; PID:
C:Comment: Prolines and lysines at the third position of the tripeptide repeating unit (C
ed and subsequently O-glycosylated.
C:Genetics:
A:Gene: GDB:COL16A1
A:CROSS-references: GDB:134045; OMIM:120326
A:Map position: lp34-lp34
A:Complex: type XVI collagen may be a homotrimer, or a heterotrimer of two alpha 1(XVI)
C:Function:
A:Description: structural component of extracellular fibrous polymer as a minor form pro
A:Note: may play a role in forming elastic connections at fibril surfaces
C:Keywords: cell binding; coiled coil; extracellular matrix; glycoprotein; hydroxylysine
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-1603/Product: collagen alpha 1(XVI) chain #status predicted <MAT>
F:22-333/Domain: amino-terminal nonhelical #status predicted <NC11>
F:334-1577/Region: interrupted helical
F:334-360/Domain: collagenous COL10 #status predicted <COL0>
F:375-505/Domain: collagenous COL9 #status predicted <COL9>
F:521-554/Domain: collagenous COL8 #status predicted <COL8>
F:539-541/Region: cell attachment (R-G-D) motif
F:572-630/Domain: collagenous COL7 #status predicted <COL7>
F:652-722/Domain: collagenous COL6 #status predicted <COL6>
F:738-875/Domain: collagenous COL5 #status predicted <COL5>
F:887-938/Domain: collagenous COL4 #status predicted <COL4>
F:973-987/Domain: collagenous COL3 #status predicted <COL3>
F:1005-1007/Region: cell attachment (R-G-D) motif

F:1011-1432/Domain: collagenous COL2 #status predicted <COL2>
F:1126-1228/Region: cell attachment (R-G-D) motif
F:1472-1577/Domain: collagenous COL1 #status predicted <COL1>
F:1578-1603/Domain: carboxyl-terminal nonhelical #status predicted <NCOL1>
F:147,327/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:		Pred. No.:	0.0076	Length:	1603
Score:			147.00	Matches:	86
Percent Similarity:			33.8%	Conservative:	13
Best Local Similarity:			22.4%	Mismatches:	112
Query Match:			3.3%	Indels:	82
Db:			2	Gaps:	17
US-10-768-158-1 (1-2419) x S23810 (1-1603)					
QY	775	AGCCCAACTCTTCCCGGCCACGGCGGCGCT---CAGCGTTGCGACCTGGTCCACC	719		
Db	1018	SerProGlyLeuProGlyProGlyLeu-ProGlyGlnArgGlyGluGluGlyProPr	1037		
QY	718	ACCTGGTGCAGTGTCCGTCAAGGCTTCCAGCTGGCGCTTGTCCAGAGACACCCCGAGG	659		
Db	1037	oGlyMetArgGlySerProGly-----ProProG1	1047		
QY	658	ATCTGGCCAGCTGTCCACCATCTCCACAGGTCCCGATGTCATCTTCATACTTGAGA	599		
Db	1047	y-----ProIleGlyProProGlyPheProGly-----	1056		
QY	598	AAAGCAGCTTCAGTCCATGCGGTCTCCAGAACTCTGCACCGTGTCAAACCCAGGAG	539		
Db	1057	-----AlaValGlySerProGlyLeuProGlyLeuGlnGlyGluArgG1	1071		
QY	538	CGGTAGC-----	521		
Db	1071	yLeuThrGlyLeuThrGlyAspLysGlyGluProGlyProGlyGlnProGlyTy-Pr	1091		
QY	520	TTCATAACCTCCGCGAGAACTCTTGAAGGTGCTCGTAGCTCATGTCGCGAGAGAG	461		
Db	1091	oGlyAlaThrGlyProProGlyLeuProGlyIleLysGlyGlu-----ArgGlyTy-Pr	1109		
QY	460	CGGTGGAAGTATGATAAGACACCAACAGATCTCTTGGGTTCCGAGCCATATAGATGACC	401		
Db	1109	rGlySerAlaGlyGluLysGlyGluProGlyProGlySerGlyGlyLeu-----Pr	1127		
QY	400	TGGAGTCTCCATTTGGAGGTCAGAGCGGAGAAAGCGGTAGCGGTCTTGATG	341		
Db	1127	oGlyProGlyProAlaGlyProArgGlyGluArgGlyProGlnGlyAsnSer-----	1145		
QY	340	AGGCGGGGAGAG-----TCAGTTCCTTGATG	314		
Db	1146	----GlyGluLysGlyAspGlnGlyPheGlnGlyProGlyPheThrGlyProThrG1	1164		
QY	313	ATGTCACGC-----CCGGCTGTGGTACTCCAGACCGGAGCTGCTCGTGGATGTC	260		
Db	1164	ySerProGlyPheProGlyLysValGlySerProGlyProGlyProGlnAlaGluLys	1184		
QY	259	ATCAAGCGATCTCATCGGGGTTCAGCGC-----CCTGGCTCACCAGTAGACCA	209		
Db	1184	sglySerGluGlyIleArgGlyProSerGlyLeuProGlySerProGlyProProGlyPr	1204		
QY	208	TCCTGCACAGCTGGT---GCCGACTT-----G	182		
Db	1204	oProGlyIleGln-GlyProAlaGlyLeuAspGlyLeuAspGlyLysAspGlyLysProG	1224		
QY	181	GGTAGGTGAGATCCACACGTCTGGCGCGCACCGGAGTTGGCGATCTCTCCATC	122		
Db	1224	lyLeuArgGlyAspProGlyProAlaGlyPro--ProGly---LeuMetGlyProProG1	1242		
QY	121	TTCCCGCGGCAGAGGGCGGCGACGCCATCGCAACTCGAAGTACTCTCTCGAAC	62		
Db	1242	yPheLysGlyLysThrGlyHisProGlyLeuProGlyProLysGlyAspCysGlyLysPr	1262		
QY	61	T---CCCCCGGGGTCTCGGGGTCTCGGCCT	34		

Db	1262	oGlyProProGlySerThrGly---ArgPro	1271		
RESULT	86				
S22917					
collagen alpha 5(IV) chain precursor, renal splice form - human					
N:Alternate names: procollagen alpha 5(IV) chain					
N:Contains: collagen alpha 5(IV) chain precursor, leukocyte splice form					
C:Species: Homo sapiens (man)					
C:Date: 30-Sep-1993 #revision 27-Feb-1997 #text_change 09-Jul-2004					
A:Accession: S22917; A54365; A57079; A37122; I54317; A34850; S18850; I56971; I76598; A35					
R:Zhou, J.; Hertz, J.M.; Leinonen, A.; Tryggvason, K.					
J. Biol. Chem. 267, 12475-12481, 1992					
A:Title: Complete amino acid sequence of the human alpha-5(IV) collagen chain and identi					
n Alport syndrome patient.					
A:Reference number: S22917; MUID:92316923; PMID:1352287					
A:Accession: S22917					
A:Molecule type: mRNA					
A:Residues: 1-967 <ZH>					
A:Cross-references: UNIPROT:P29400; UNIPARC:UPI0000173BDF; GB:M90464; NID:gl80826; PIDN:					
R:Zhou, J.; Leinonen, A.; Tryggvason, K.					
J. Biol. Chem. 269, 6608-6614, 1994					
A:Title: Structure of the human type IV collagen COL4A5 gene.					
A:Reference number: A54365; MUID:94165049; PMID:8120014					
A:Accession: A54365					
A:Molecule type: DNA					
A:Residues: 1-922 <ZH2>					
A:Cross-references: UNIPARC:UPI0000173BE0; GB:U04470; NID:g463378; GB:U04520; NID:g46342					
R:Zhou, J.; Mochizuki, T.; Smeets, H.; Antignac, C.; Laurila, P.; de Paape, A.; Tryggvas					
Science 261, 1167-1169, 1993					
A:Title: Deletion of the paired alphas(IV) and alpha6(IV) collagen genes in inherited sm					
A:Reference number: A57079; MUID:93361972; PMID:8356449					
A:Accession: A57079					
A:Molecule type: DNA					
A:Residues: 1-27 <ZH4>					
A:Cross-references: UNIPARC:UPI000007378A; GB:Z37153; NID:g587203; PIDN:CAA85512.1; PID					
R:Pihlajaniemi, T.; Pohjolainen, E.R.; Myers, J.C.					
J. Biol. Chem. 265, 13758-13766, 1990					
A:Title: Complete primary structure of the triple-helical region and the carboxyl-termin					
A:Reference number: A37122; MUID:90337990; PMID:2380186					
A:Accession: A37122					
A:Molecule type: mRNA					
A:Residues: 84-439, 'GS', 442-624, 'LALQ', 629-666, 'FR', 669-887, 'R', 889-1264, 1271-1691 <PTH>					
A:Cross-references: UNIPARC:UPI0000173BE1; GB:J05558; EMBL:M58526; NID:gl314209					
A:Note: submitted to the EMBL Data Library, February 1991					
R:Renier, A.; Seri, M.; Myers, J.C.; Pihlajaniemi, T.; Massella, L.; Rizzoni, G.; De Ma					
Hum. Mol. Genet. 1, 127-129, 1992					
A:Title: De novo mutation in the COL4A5 gene converting glycine 325 to glutamic acid in					
A:Reference number: I54317; MUID:93244772; PMID:1363780					
A:Accession: I54317					
A>Status: preliminary; translated from GB/EMBL/DBDJ					
A:Molecule type: mRNA					
A:Residues: 313-324, 'E', 326-330 <REN>					
A:Cross-references: UNIPARC:UPI000016B3D0; GB:S59334; NID:g299946; PIDN:AAD13909.1; PID:					
R:Hostikka, S.L.; Eddy, R.L.; Byers, M.G.; Hoeythya, M.; Shows, T.B.; Tryggvason, K.					
Proc. Natl. Acad. Sci. U.S.A. 87, 1606-1610, 1990					
A:Title: Identification of a distinct type IV collagen alpha chain with restricted kidne					
A:Reference number: A34850; MUID:90160375; PMID:1689491					
A:Accession: A34850					
A:Molecule type: mRNA					
A:Residues: 924-1264, 1271-1691 <ZH3>					
A:Cross-references: UNIPARC:UPI000016A437; EMBL:M63456; EMBL:M63457; EMBL:M					
R:Zhou, J.; Hostikka, S.L.; Chow, L.T.; Tryggvason, K.					
Genomics 9, 1-9, 1991					
A:Title: Characterization of the 3' half of the human type IV collagen alpha-5 gene that					
A:Reference number: A37969; MUID:91169491; PMID:2004755					
A:Accession: S18850					
A:Molecule type: DNA					
A:Residues: 924-1264, 1271-1691 <ZH3>					
A:Cross-references: UNIPARC:UPI000016A437; EMBL:M63456; EMBL:M63457; EMBL:M					
R:Guo, C.; Van Damme, B.; Van Damme-Lombaerte, R.; Van den Berghe, H.; Cassiman, J.J.; M					

Kidney Int. 44, 1316-1321, 1993
A:Title: Differential splicing of COL4A5 mRNA in kidney and white blood cells: a complex
A:Reference number: 156971; PMID:94133540; PMID:8301933
A:Accession: 156971
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1258-1276 <GU01>
A:Cross-references: UNIPARC:UPI000016B421; GB:S69168; NID:g545095; PIDN:AAC60612.1; PID:
A>Note: kidney splice form
A:Accession: 176598
A>Status: translated from GB/EMBL/DBJ
A:Residues: 1284-1291 'TFLGYLACLV' <GU02>
A:Cross-references: UNIPARC:UPI000011DDFD; GB:S69169; NID:g545097; PIDN:AAC60613.1; PID:
A>Note: frameshift mutation in patient with Alport syndrome
R:Myers, J.C.; Jones, T.A.; Pohjola, E.R.; Kadr, A.S.; Goddard, A.D.; Sheer, D.; So
Am. J. Hum. Genet. 46, 1024-1033, 1990
A:Title: Molecular cloning of alpha5(IV) collagen and assignment of the gene to the regi
A:Reference number: A35335; MUID:90252791; PMID:2339699
A:Accession: A35335
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1448-1477 <MYE>
A:Cross-references: UNIPARC:UPI00001738E2
R:Nakazato, H.; Hattori, S.; Ushijima, T.; Matsuura, T.; Koitabaehi, Y.; Takada, T.; Yoe
Kidney Int. 46, 1307-1314, 1994
A:Title: Mutations in the COL4A5 gene in Alport syndrome: a possible mutation in primor
A:Reference number: 156975; MUID:95156893; PMID:7853788
A:Accession: 156975
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1595-1602 <NAK>
A:Cross-references: UNIPARC:UPI00000004F8; GB:S75903; NID:g913882; PIDN:AAB33374.1; PID:
A>Note: premature termination mutation from a patient with Alport syndrome; one other mu
R:Lemnick, H.H.; Schroeder, C.H.; Brunner, H.G.; Nelen, M.R.; Zhou, J.; Tryggvason, K.;
Genomics 17, 485-489, 1993
A:Title: Identification of four novel mutations in the COL4A5 gene of patients with Alpo
A:Reference number: 154188; MUID:94010948; PMID:8406498
A:Accession: 154188
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1604-1607 'VHDAYKC' <LEW>
A:Cross-references: UNIPARC:UPI000011F85C; GB:S65767; NID:g425563; PIDN:AAD13967.1; PID:
A>Note: frameshift mutation from a patient with Alport syndrome; five other mutations a
C:Comment: Prolines and lysines at the third position of the tripeptide repeating unit
ed and subsequently O-glycosylated.
C:Genetics:
A:Gene: GDB:COL4A5; ATS
A:Cross-references: GDB:120596; OMIM:303630
A:Map position: Xq22-Xq22
A:Introns: 27/3; 47/3; 77/3; 92/3; 107/3; 128/3; 146/3; 155/3; 182/3; 203/3; 215/3; 229/
/3; 799/1; 837/1; 893/1; 923/1; 1006/1; 1036/1; 1082/3; 1125/1; 1152/1; 1185/1; 1
A>Note: the alpha 5(IV) and alpha 6(IV) chain genes are encoded on opposite strands with
C:Complex: This minor type IV collagen is thought to form a heterotrimer of two alpha 5
among trimer amino-terminal domains (with disulfide and desmosine cross-links), dimeric a
er associations in the interrupted helical domain (with disulfide and desmosine cross-li
C:Function:
A:Description: minor structural component of extracellular basement membrane
A:Superfamily: collagen alpha 1(IV) chain
C:Keywords: Alport syndrome; basement membrane; coiled coil; extracellular matrix; glyco
F:1-26/Domain: signal sequence #status predicted <SIG>
F:27-1691/Product: collagen alpha 5(IV) chain, renal splice form #status predicted <MAT1
F:27-1264,1271-1691/Product: collagen alpha 5(IV) chain, leukocyte splice form #status P
F:27-41/Domain: amino-terminal nonhelical, NC2 #status predicted <NC2>
F:42-1462/Region: interrupted helical
F:1463-1691/Domain: carboxyl-terminal nonhelical, NC1 #status predicted <NC1>
F:1473-1573/Domain: collagen IV carboxyl-terminal repeat <CT1>
F:1583-1687/Domain: collagen IV carboxyl-terminal repeat <CT2>
F:29-32,38,40,124,451,481,484/Disulfide bonds: interchain #status predicted
F:1482-1570/Disulfide bonds: (covalent) #status predicted
F:1482-1570,1515-1573/Disulfide bonds: (or 1482-1573, 1515-1570) #status predicted
F:1527-1533,1638-1644/Disulfide bonds: #status predicted
F:1592-1684,1626-1687/Disulfide bonds: (or 1592-1687, 1626-1684) #status predicted

Alignment Scores:
Pred. No.: 0.00769 Length: 1691
Score: 147.00 Matches: 75
Percent Similarity: 35.8% Conservative: 22
Best Local Similarity: 27.7% Mismatches: 108
Query Match: 3.3% Indels: 66
DB: 1 Gaps: 16

US-10-768-158-1 (1-2419) x S22917 (1-1691)

QY 765 TTCCTCCGCGCCACGGCAGCGCTTCAGCGTTGCAGCACTGGTCCACGAGCTGGTGGCAGT 706
|||||
Db 401 PheProGlyGluAsgGly-----GlnLysGlyAspGluGlyProPro---GlyLeuSer 417
QY 705 GCTCCG-----TCAGGGCTTCACCTGGGCGCTTGTTCACAGAGACACCCCA 661
|||||
Db 418 IleProGlyProProGlyLeuAspGlyGlnProGlyAlaProGlyLeuProGlyProPro 437
QY 660 GGAATCTGGCCAGCTGCTCCACCATCGTCACCAAGTCCCGATCGCTTTCATCTTGA 601
|||||
Db 438 Gly-----ProAlaGlyProHisIleProProSerAspGluIleCys----- 451
QY 600 GAAAAAGCACCTTCGAGTCCATGCGTCTCCAGAACTCTCGCAGCTGCTCAAAACCA 541
|||||
Db 452 -----GluProGlyProProGlyProProGlySerProGlyAspLysGlyLeuGln 468
QY 540 AGCGGTACCCAGCTTATCATTTCAATAAACCTCCCGGAGAAATCTTGAAGGTGCTCGGT 481
|||||
Db 469 GlyGluGlnGlyValLysGlyAsp-----LysGlyAspThrCysPheAsnCysIleGly 486
QY 480 AGCTCATGTCCTCCGAGAGAGCGGTGGAAGTAAAGACACCAACCAAGTCTTGGGGT 421
|||||
Db 487 ThrGlyIleSerGlyProProGlyGlnProGlyLeuProGlyLeuProGlyPro----- 504
QY 420 TCGAGCCATATAGATGACCTTGGAGTCTCCATTTGTGGAGGTTCAGAGGCGAGAACGCGT 361
|||||
Db 505 -----ProGlySerLeuGlyPheProGlyGlnLysGlyLysGly 518
QY 360 AGGCAGGTGCTCTTGATAGAGGGGAGAGTTCAGTTCTTGTATGATGTCAGGCGCG 301
|||||
Db 519 GlnAlaGlyAlaThr-----GlyProLysGlyLeuPro---GlyIleProGly--- 533
QY 300 GCTGTGGGTACTCCAGGA-----CCGGGAGTCTCGTCGATGAT 262
|||||
Db 534 AlaProGlyAlaProGlyPheProGlySerLysGlyGluProGlyAspIleLeuThrPhe 553
QY 261 TCATCAAGCCGATCTCATCGGGGTGAG-----CGCCCTGGCTCACCAAGTAGACCA--- 211
|||||
Db 554 ProGlyMetLysGlyAspLysGlyGlnLeuGlySerProGlyAlaProGlyLeuProGly 573
QY 210 -----CCTCTGAGCAAGCTGGTGGCGGACTTGGCTCGAACTCCCGGGGTGCTGGGGTCT 160
|||||
Db 574 LeuProGlyThrProGlyGlnAspGlyLeu-----ProGlyLeu 586
QY 159 CGCTGGGCGGACCGGGAGTGGCGATCTCTCCATCTTCCGCGGCGAGAGGCGGCA 100
|||||
Db 586 uProGlyProLysGlyGluProGlyGlyLe-ThrPhe-----LysGlyGluA 602
QY 99 GCGGCACGCCCTGGAACCTCGAAGTACTTGTCTCGAACTCCCGGGGTGCTGGGGTCT 40
|||||
Db 602 rgGlyProProGlyAsn-----ProGlyLeuProGlyLeuP 614
QY 39 CGGCTCGCTCTCCGCCATGCGCGCGCG 11
|||||
Db 614 roGlyAsnIleGlyProMetGlyProPro 623

RESULT 87

CSB01S
Collagen alpha 1(I) chain - bovine (tentative sequence) (fragments)
C:Species: Bos primigenius taurus (cattle)
C>Date: 24-Apr-1984 #sequence revision 31-Dec-1993 #text change 09-Jul-2004
C:Accession: A91193; A91229; A91387; A91201; A91200; A43048; A02853

R;Rauterberg, J.; Timpl, R.; Furthmayr, H.
 Eur. J. Biochem. 27, 231-237, 1972
 A:Title: Structural characterization of N-terminal antigenic determinants in calf and hu
 A:Reference number: A91193; MUID:72255334; PMID:4115172
 A:Accession: A91193
 A:Molecule type: protein
 A:Residues: 1-19 <RAU>
 A:CROSS-references: UNIPROT:P02453; UNIPARC:UPI0000173B51
 A:Experimental source: skin
 A:Note: the epsilon carbon of Lys-9, by homology with the rat alpha 1(I) chain, is conve
 R;Fietzek, P.P.; Kuehn, K.
 Eur. J. Biochem. 52, 77-82, 1975
 A:Title: The covalent structure of collagen: amino-acid sequence of the cyanogen-bromide
 A:Reference number: A91229; MUID:76022320; PMID:1164916
 A:Accession: A91229
 A:Molecule type: protein
 A:Residues: 20-145 <FIE>
 A:CROSS-references: UNIPARC:UPI0000173B52
 A:Experimental source: skin
 A:Note: Lys-103 is hydroxylated and binds glucosylgalactose
 R;Fietzek, P.P.; Wendt, P.; Kell, I.; Kuehn, K.
 FEBS Lett. 26, 74-76, 1972
 A:Title: The covalent structure of collagen: amino acid sequence of alpha1-CB3 from calf
 A:Reference number: A91387; MUID:73049499; PMID:4673951
 A:Accession: A91387
 A:Molecule type: protein
 A:Residues: 146-294 <F12>
 A:CROSS-references: UNIPARC:UPI0000173B53
 A:Experimental source: skin
 R;Fietzek, P.P.; Rexrodt, F.W.; Hopper, K.E.; Kuehn, K.
 Eur. J. Biochem. 38, 396-400, 1973
 A:Title: The covalent structure of collagen. 2. The amino-acid sequence of alpha1-CB7 fr
 A:Reference number: A91211; MUID:74086118; PMID:4359390
 A:Accession: A91211
 A:Molecule type: protein
 A:Residues: 295-562 <F13>
 A:CROSS-references: UNIPARC:UPI0000173B54
 A:Experimental source: skin
 R;Wendt, P.; Mark, K.V.D.; Rexrodt, F.; Kuehn, K.
 Eur. J. Biochem. 30, 169-183, 1972
 A:Title: The covalent structure of collagen. The amino-acid sequence of the 112 residues
 A:Reference number: A91201; MUID:73042276; PMID:4343808
 A:Accession: A91201
 A:Molecule type: protein
 A:Residues: 563-675 <WEN>
 A:CROSS-references: UNIPARC:UPI0000173B55
 A:Experimental source: skin
 R;Fietzek, P.P.; Rexrodt, F.W.; Wendt, P.; Stark, M.; Kuehn, K.
 Eur. J. Biochem. 30, 163-168, 1972
 A:Title: The covalent structure of collagen. Amino acid sequence of peptide alpha1-CB6-C
 A:Reference number: A91200; MUID:73042275; PMID:4343807
 A:Accession: A91200
 A:Molecule type: protein
 A:Residues: 676-758 <F14>
 A:CROSS-references: UNIPARC:UPI0000173B56
 A:Experimental source: skin
 A:Note: Pro-726 is the only 3-hydroxyproline and the only hydroxylated proline in positi
 R;Rauterberg, J.; Fietzek, P.; Rexrodt, F.; Becker, U.; Stark, M.; Kuehn, K.
 FEBS Lett. 21, 75-79, 1972
 A:Title: The amino acid sequence of the carboxyterminal nonhelical cross link region of
 A:Reference number: A43048
 A:Accession: A43048
 A:Molecule type: protein
 A:Residues: 759-779 <RA2>
 A:CROSS-references: UNIPARC:UPI0000173B57
 A:Experimental source: skin
 C:Comment: Lysines at positions 115, 124, 274, 346, 424, 496, 658, and 670 may be hydrox
 C:Comment: Prolines in the third position of the tripeptide repeating unit (G-X-Y) are h
 C:Comment: The order of the eight CNBr peptides in the alpha 1(I) chain of bovine skin d
 9, 149, 268, and 217 residues.
 C:Comment: The complete chain contains 1052 residues.
 C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;
 C:Keywords: coiled coil; extracellular matrix; glycoprotein; pyroglutamic acid; trimer;

F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Alignment Scores:

Pred. No.: 0.00709 Length: 779
 Score: 146.50 Matches: 83
 Percent Similarity: 29.9% Conservative: 13
 Best Local Similarity: 25.9% Mismatches: 112
 Query Match: 3.3% Indels: 113
 DB: 1 Gaps: 19

US-10-768-158-1 (1-2419) x CGB01S (1-779)

QY	783	CCTTCCACAGCCCAACTCTTCCCGGCCCA-----CGGGCAGGCGCTCAGCGTTCCAGC	730
DB	432	ProProGlyAlaThrGlyPheProGlyAlaAlaGlyArgValGlyProPro-----	448
QY	729	ACTGGTCCACACAGCTGGTGGCAGTGTCCGTCCAGGCTTCCAGCT-----	685
DB	449	-----GlyProSerGlyAsnAlaGlyProProGlyProGlyProAlaGlyLysGlu	466
QY	684	-----GGCCTTGTCCACAGGACCCCCAGGAATCTGGCCA-----GCTGCTCCA	640
DB	467	GlySerLysGlyProArgGlyGluThrGlyProAlaGlyArgProGlyGluValGlyPro	486
QY	639	CCATCGTCCACAGCTCCGATGCATGTCTTCACTTGAGAAAA-----GCACGT	589
DB	487	ProGlyProProGlyProAlaGly-----GluLysGlyAlaProGlyAla	501
QY	588	TCGAGTCCATGCGGTCTCCAGAACTCTCGACGTCTCAACACGAGGACCGTAGCCCA	529
DB	502	AspGlyProAlaGlyAlaProGlyThrProGlyProGlnGlyIleAlaGlyGlnArgGly	521
QY	528	GCTTATCATTTCAAAACCTCCCGCAGAAATCTTGAAGGTGCTC-----GGT	481
DB	522	ValValGly-----LeuProGlyGlnArgGlyGluArgGlyPheProGlyLeuProGly	539
QY	480	ACCTCATGG-----TCCGAGAGACGGTGGAACTGTAATAAGACACACCA	433
DB	540	ProSerGlyGluProGlyLysGlnGlyProSerGlyAlaSerGlyGluArgGlyProPro	559
QY	432	GATCCTTGGGTTCGAGCCATATAGATGACCTTCGAGTCTCCATTGTGAGGTCCAGAG	373
DB	560	GlyProMetGly-----ProProGlyLeu-----AlaGlyProPro	571
QY	372	GCAGAAAGCGGTAGGCGAGTGGCTCTTGATGAGCGGGGAGAGGTCACTTCTTGTATGA	313
DB	572	GlyGluSerGlyArgGluGly-----	578
QY	312	TGTCAGGCGCGGTGGTGGTACTCCAGGA-----CCGGAGCTGCTGCTCGA	265
DB	579	AlaProGly--AlaGluGlySerProGlyArgAspGlySerProGlyAla-----	594
QY	264	TGTTTCATCAAGCCGATCTCATCGGGGTCCAGCGC-----CCT	229
DB	595	-----LysGlyAspArgGlyGluThrGlyProAlaGlyAlaProGlyProPro	610
QY	228	GGCTCACAAGTAGACCACTCTCTCCAGCAAGCTGGTCCGAGACTTGGGTAGTGACGA	169
DB	611	GlyAlaProGlyAlaProGlyProAlaGlyLysSer-----	625
QY	168	TCCACAGCTCGTGGCGCCGACCGGAAAGTTGGCGATCTCTCCATCTTCCCGCGGCGA	109
DB	626	GlyAspArgGlyGluThrGlyProAlaGlyProLleGlyProValGlyProAlaGlyAla	645
QY	108	AGGGCGGCGAGCGCAGCATCGAACTCTGCTCTCGAACT-----	61
DB	646	ArgGlyProAlaGlyProGlnGlyProArgGly***LysGly***ThrGly***Gly	665
QY	60	-----CCCCCG	55
DB	666	***ArgGlyIleLysGlyHisArgGlyPheSerGlyLeuGlnGlyProProGlyProPro	685
QY	54	GGGTGCTGGGGTCTCGG-----CCTGCTCTCGG	25

C;Comment: Type II collagen molecules are trimers of identical alpha 1(II) chains, generated the CNBr peptides was determined as 1-4-2-3-6-12-11-8-10-5-9-7-14-15.

C;Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology; C;Keywords: cartilage; coiled coil; extracellular matrix; glycoprotein; hydroxyllysine; F;493-673/Domain: carboxyl-terminal propeptide (fragment) #status predicted <CTP> F;493-673/Domain: fibrillar collagen propeptide (fragment) #status predicted <FC> F;9,102,114,123,189,423,435/Modified site: 5-hydroxyllysine (Lys) #status experimental F;9,102,114,123,189,423,435/Binding site: carboxylate (Lys) (covalent) #status experimental F;574/Binding site: carboxylate (Asn) (covalent) #status predicted

Alignment Scores:
Pred. No.: 0.0093 Length: 673
Score: 145.00 Matches: 78
Percent Similarity: 33.6% Conservative: 21
Best Local Similarity: 26.4% Mismatches: 130
Query Match: 3.3% Indels: 66
DB: 1 Gaps: 15

US-10-768-158-1 (1-2419) x CGB06C (1-673)

QY 759 GCGCCAGCGGAGGCGCTTCAGCGTTCAGCACTGGTCCACAGCTGGTGGCGAGTCTCG 700

Db 22 GlyProAargGlyProPro-----GlyProAlaGlyAlaProGlyPro 35

QY 699 TCAGGGCTCCAGCTGGGCGTGTGCACAGGACACCCCGAGGAATC----- 655

Db 36 GInGlyPheGlnGlyAsnProGlyGluProGlyGluProGlyValSerGlyProMetGly 55

QY 654 -----TGCCAGCTGCTCCACATCGTCACCGAGTCCCGATGTCATGCTT----- 610

Db 56 ProArgGlyProProGlyProProGlyLysProGlyAspGlyGluAlaGlyLysPro 75

QY 609 ---CATACTGTGAAGAACACAGTTCAGTCCATCGGTGCT-----CCAGAACT 562

Db 76 GlyLysSerGlyGluArgGlyProProGlyProGlnGlyAlaArgGlyPheProGlyThr 95

QY 561 CTGACAGTGTCAAAACAGGAGCGGTAGCCAGCTTATCATTCATAAACCTCCGCGAGA 502

Db 96 ProGlyLeuProGlyValLysGly-----HisArgGlyTyrProGly--- 109

QY 501 ATTCTTGAAGGTGCTCCGTAGCTCATGGTCC-----GCAGAGAGCGGTGA 454

Db 110 ---LeuAspGlyAlaLysGlyGluAlaGlyAlaProGlyValLysGlySerGlySer 128

QY 453 ACTGATAATAGACACACAGATCTTGGGTGGCGAGCCATATAGATGACTTGGAGT 394

Db 129 ProGly****GlySerProGlyProMetGly-----ProArgGly 142

QY 393 CTCCATTGTGGAGTTCAGAGGCGAGAACCGGTAGGCGAGGTGGCTTTGATGAGCGGG 334

Db 143 Leu-----ProGlyGluArgGlyArgThrGlyProAlaGlyAlaAlaGlyAlaArgGly 160

QY 333 GAGAGTTCAGTCTTGTGATGATG-----CCAGGCGCG----- 301

Db 161 AsnAspGlyGlnProGlyProAlaGlyProProGlyProValGlyProAlaGlyGlyPro 180

QY 300 -----GCTGTGGTACTTCCAGGACCGGAGCTCTCGTTCGATGT 262

Db 181 GlyPheProGlyAlaProGlyAlaLysGlyGluAlaGlyProGlnGlyAlaArgGlyPro 200

QY 261 TCATCAAGCGGATCTCATCGGGTTCAGCGCTTCAGGCTACCAAGTAGACCACTCTCGCA 202

Db 201 GluGlyAlaGlnGlyProArgGlyGlu---ProGlyThrProGlyAlaProGlyProAla 219

QY 201 GCAAGTGTGCGGACTTGGGTAGGTGACGATCCACAGTCCAGTCCGCGCGCGAGCGGA 142

Db 220 Gly-AlaAlaGlyAsnProGlyAlaAspGlyLysPro-GlyAlaLysGlySerAlaGlyA 239

QY 141 AGTTGGCATCTCTCCATC-----TTCCGCGGCGAGAGGCG-----G 103

Db 239 laProGlyIleAlaGlyAlaProGlyPheProGlyAlaArgGlyProProGlyProThrG 259

QY 102 GCAGCGCGACCGCACTCGAAGTACTTGTCTCGAAGTCTCCCGCGGGTCTCGGGG 43

Db 259 lyAlaSerGlyProLeuGlyProLysGly---GlnThrGlyGluProGlyIleAlaGlyP 278

QY 42 TCTCGGCTCCTCTCCGCCATGCGCGCGCGTCCGCTCG 2

Db 278 heLysGlyGluGlnGlyProLysGlyGluProGlyProAla 291

RESULT 91

T45467

collagen alpha 1(II) chain precursor [imported] - horse

N;Alternate names: type II collagen

C;Species: Equus caballus (domestic horse)

C;Date: 31-Jan-2000 #sequence_revision 31-Jan-2000 #text_change 09-Jul-2004

C;Accession: T45467

R;Richardson, D.W.; Dodge, G.R.

submitted to the EMBL Data Library, June 1996

A;Description: Cloning of equine type II collagen and modulation of its expression in e

A;Reference number: Z22977

A;Accession: T45467

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-1418 <RIC>

A;Cross-references: UNIPROT:Q28396; UNIPARC:UPI00008834A; EMBL:U62528; PIDN:AA05773.1

C;Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;

Alignment Scores:
Pred. No.: 0.0105 Length: 1418
Score: 145.00 Matches: 78
Percent Similarity: 33.6% Conservative: 21
Best Local Similarity: 26.4% Mismatches: 130
Query Match: 3.3% Indels: 66
DB: 2 Gaps: 15

US-10-768-158-1 (1-2419) x T45467 (1-1418)

QY 759 GCGCCAGCGGAGGCGCTTCAGCGTTCAGCACTGGTCCACAGCTGGTGGCGAGTCTCG 700

Db 138 GlyProAargGlyProPro-----GlyProAlaGlyAlaProGlyPro 151

QY 699 TCAGGGCTTCAGCTGGGCGTGTGCACAGGACACCCCGAGGAATC----- 655

Db 152 GInGlyPheGlnGlyAsnProGlyGluProGlyGluProGlyValSerGlyProMetGly 171

QY 654 -----TGCCAGCTGCTCCACATCGTCACCGAGTCCCGATGTCATGCTT----- 610

Db 172 ProArgGlyProProGlyProProGlyLysProGlyAspGlyGluAlaGlyLysPro 191

QY 609 ---CATACTTGAAGGTGCTTCGAGTCCATCGGTGCT-----CCAGAACT 562

Db 192 GlyLysSerGlyGluArgGlyProProGlyProGlnGlyAlaArgGlyPheProGlyThr 211

QY 561 CTGACAGTGTCAAAACAGGAGCGGTAGCCAGCTTATCATTCATAAACCTCCGCGAGA 502

Db 212 ProGlyLeuProGlyValLysGly-----HisArgGlyTyrProGly--- 225

QY 501 ATTCTTGAAGGTGCTTCGAGTCCATCGGTGCT-----GCAGAGAGCGGTGA 454

Db 226 ---LeuAspGlyAlaLysGlyGluAlaGlyAlaProGlyValLysGlySerGlySer 244

QY 453 ACTGATAATAGACACACAGATCTTGGGTGGCGAGCCATATAGATGACTTGGAGT 394

Db 245 ProGlyGluAsnGlySerProGlyProMetGly-----ProArgGly 258

QY 393 CTCCATTGTGGAGTTCAGAGGCGAGAACCGGTAGGCGAGGTGGCTTTGATGAGCGGG 334

Db 259 Leu-----ProGlyGluArgGlyArgThrGlyProAlaGlyAlaAlaGlyAlaArgGly 276

QY 333 GAGAGTTCAGTCTTGTGATGATG-----CCAGGCGCG----- 301

Db 277 AsnAspGlyGlnProGlyProAlaGlyProProGlyProValGlyProAlaGlyGlyPro 296

QY 300 -----GCTGTGGTACTCCAGGACCGGAGTCTCGTTCGATGT 262

Db 297 GlyPheProGlyAlaProGlyAlaLysGlyGluAlaGlyProThrGlyAlaArgGlyPro 316
Qy 261 TCATCAAGCGATCTCATCGGGTCAGCGCCCTGGCTCACCAGTAGACCACTCTCGCA 202
Db 317 GluGlyAlaGlnGlyProArgGlyGlu---ProGlyThrProGlySerProGlyProAla 335
Qy 201 GCAAGCTGTGCGGAGCTGGGGTAGGTGACGATCCACACGTCGCTGGGCGGACCGGA 142
Db 336 Gly-AlaAlaGlyAsnProGlyThrAspGlyIlePro-GlyAlaLysGlySerAlaGlyA 355
Qy 141 AGTTGGCGATCTCTCCATC-----TTCGCGCGGAGAGGCG-----G 103
Db 355 laProGlyIleAlaGlyAlaProGlyPheProGlyProArgGlyProProGlyProGlnG 375
Qy 102 GCAGCGCAGCGCATCGAACTCGAAGTACTTGTCTCGAACTCCCGCGGGTGTGGGG 43
Db 375 lYAlaThrGlyProLeuGlyProLysGly--GlnThrGlyGluProGlyIleAlaGlyP 394
Qy 42 TCTCGGCTCGCTCTCCGCATGCCCGCGCGCTGCGCGTCG 2
Db 394 heLysGlyGluGlnGlyProLysGlyGluProGlyProAla 407
RESULT 92
T19732
hypothetical protein C34F6.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T19732
R:White, S.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z19171
A:Accession: T19732
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-283 <WIL>
A:Cross-references: UNIPROT:O17642; UNIPARC:UPI0000082853; EMBL:Z81479; PIDN:CAB03942.1;
A:Experimental source: clone C34F6
C:Genetics:
A:Gene: CESP:C34F6.3
A:Map position: X
A:Introns: 64/2
Alignment Scores:
Pred. No.: 0.00807 Length: 283
Score: 144.50 Matches: 78
Percent Similarity: 29.4% Conservative: 19
Best Local Similarity: 23.6% Mismatches: 93
Query Match: 3.3% Indels: 140
DB: 2 Gaps: 17
US-10-768-158-1 (1-2419) x T19732 (1-283)
Qy 841 CCCATCTCTGTTTATACACCAAGTCAAACTTCTCATTCATGAGACG----- 794
Db 36 ProLeuPheCysValHisMetGlnSerValThrSerGlyLeuSerGluGluLeuLeuPhe 55
Qy 793 -----GTGAAGATGTCCTTCCACAGCCCAACTCTTCCCGGCGCC 755
Db 56 CysLysSerLysAsnValTyrIleLysGlyGluLeuGlnLeuSerValThrArgGlu 75
Qy 754 ACGGCGAGCGCTCAGCGT-----TGCAGCACTGGTCCACCACTCGG 713
Db 76 AlaGlyArg-GlnLysArgGlnThrProGlnThrCysCysSerCysGly-----IleG 93
Qy 712 TGGCACTGCTCGTCAAGGCTTCCAGCTGCGCTTGTCTACAGACACCCCGAGGAATCTG 653
Db 93 yGluThrGlyProAlaGlyValProGlyGln-----GluGlyAlaProGlyAsnAspG 111
Qy 652 GCCAGCTGCTCCACCATCTGCACGAGTCCCGATGCA----- 616
Db 111 yLysAlaGlyGlnProGlyAlaAspAlaAspGluGlnGlyPheHisTyrLy 131
Qy 615 -----TGCTTTCATCTTGAGAAAAAGCACGTTCCAGTCCATGCGGTCTCC 569

Db 131 sAlaProGluPheCysPhe-----AspCysProAlaGlyProPr 144
Qy 568 CAGAACTCTGTGCACGTGCTCAAAACAGGAGCGTAGCCAGCTTATCATTAACCTC 509
Db 144 o----- 144
Qy 508 CGCGAGAATTTTGAAGGTGCTCGTAGCTCATGGTCCGAGAGAGCGGTGGAACCTGA 449
Db 145 -GlyAlaVal-----GlyGlyProG 151
Qy 448 TAATAAGACACACACAGATCTTGGGTTGGAGCCATATAGATGACCTTGGAGTCCA 389
Db 151 yProLysGlyProProGlyProProGly----- 160
Qy 388 TTGTGGAGGTGAGAGGCGAGAAAGGTAGGCGAGGTGGCTCTTGATGAGGGGGGAGAG 329
Db 161 -----GlyProGlyGluLeuGlyGlyProGlyArgGlyGlyAsnAr 174
Qy 328 GTCACTTCTTGATGATGTCCAGGCCCGGCTGTGGGTACTCCAGGA----- 283
Db 174 ggLyPro-----ProGlyProArg---GlyProProGlyGluAlaGlyProAs 189
Qy 282 -----CCGGGAGCTGCTCGTCGATGTTTCATCAAGCCGATCTCATCGGG 239
Db 189 pGlyGluGlyArgProGlyGlnAlaGlyGlnThrArgSerAlaProSerProProG 209
Qy 238 TCAG-----CGCCCTGGCTCACCAAGTAGACCACTCTCTGCAACGAAGCTG 194
Db 209 yGlnProGlyGlnProGlyGluProGlySerProGlyGluProGlyProAsp----- 226
Qy 193 CTGCGGACTTGGGTAGGTGACATCCACACGTCGCTGGGCGCCGCGGAAAGTTGGCG 134
Db 227 -----GlyArgAlaGlyHisProGlyArgAsnGlyProProGly----- 239
Qy 133 ATCTCTCTCATCTTCCCG----- 110
Db 240 ----ProProGlyAspAsnGlyGlyGlnGlyGluProGlyLysAspGlyGluAspGly 258
Qy 109 AAGGCGCGCAGCGCCGACCATGGAACCTCGAAGTACTTGTCTCTCGAATCTCCCGGGGTG 50
Db 258 uAsnGlyAlaAlaGlyAlaAlaGlyProLysGlySerCysAspHisCysProPro----- 276
Qy 49 CTGGGGGTCTCGGCTCGCTCTCCGCCA 22
Db 277 -----ProArgThrAlaPro 281
RESULT 93
S34665
collagen, cuticular - root-knot nematode (Meloidogyne incognita)
C:Species: Meloidogyne incognita
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C:Accession: S34665
R:van der Eycken, W.V.; de Almeida Engler, J.; van Montagu, M.; Gheysen, G.
submitted to the EMBL Data Library, July 1993
A:Description: Identification and analysis of a cuticular collagen gene from the plant-
A:Reference number: S34665
A:Accession: S34665
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-286 <VAN>
A:Cross-references: UNIPROT:Q25467; UNIPARC:UPI000016BF3F; EMBL:Z24734; NID:g395144; PFI
Alignment Scores:
Pred. No.: 0.00809 Length: 286
Score: 144.50 Matches: 64
Percent Similarity: 34.2% Conservative: 17
Best Local Similarity: 27.0% Mismatches: 97
Query Match: 3.3% Indels: 59
DB: 2 Gaps: 11
US-10-768-158-1 (1-2419) x S34665 (1-286)

```
QY 735 TGCAGCAGCTGGTCCACAGCTGGTGGCAGTGTCTCCAGGCTTCCAGCTGGGCTTGT 676
Db 80 CysAlaProGlyProPro-----GlyProProGlyProProGlyGlnProGly 95
QY 675 CACGAGCACCCCGAGGAATCGGCAGCTGCTCCACCATCCTCCACAGGTCCTCCGATGCA 616
Db 96 HisProGlyGlnProGly-----HisValGlyGlnProGlySerProGlyGlnProAla 113
QY 615 -----TGCTCTCATCTTGAGAAAGACAGCTTCAGTCCATGGGTGCTCCAGAACT 562
Db 114 ProProCysProLeuProGlnGlnAlaCysGlnArgCysProAlaGlyAlaProGlyThr 133
QY 561 CCGTCAGCTGCTCABACAGGAGCGTAGCCAGCTTATCATTCATAAACTCCGCGCAGA 502
Db 134 ProGlyLys-----GlnGlyProAlaGlyGlnProGlyGlnProGlyArg 148
QY 501 ATTCTTGAAGTGCTCGTAGCTCATGCTCCGACAGAGCGGTGGAATCATATAAAG 442
Db 149 -----ProGlyAlaProGlyLysSerSerGlyAla 158
QY 441 ACACACAGATCCTTGGGGTTGGAGCCATATAGATGACCTTGAGAGTCTCCATTGTGGA 382
Db 159 GlyProProGlyProAlaGlyProGlnGlyProProGlyProAlaGlyLysHisGlyGly 178
QY 381 ---GGTCAGAGGCGAAGCGTAGGCGAGGTGGCTCTTGATGAGCGGGGAGAGTCA 325
Db 179 ProGlyGlnProGlyGlnProGlyLysAsnGlyValSer-----HisProThr 194
QY 324 GTTCTTGATGATGATCCAGCGCGCTGTGGTACTTCCAGGACCGGAGCTCTCGTCTGA 265
Db 195 IlePro-----GlyProLysGlyProSerGlySer----- 204
QY 264 TTTTCATCAAGCCGATCTCATCGGGTTCAGCGGCTCAGCCGCTCGCTACCAAGTAGACCACTCT 205
Db 205 -----ProGlyGlnProGlyLysProGlyPro 213
QY 204 GCAGCAAGCTGTGCGGACTTGG-----GGTAGGTGACGA 169
Db 214 AlaGlyValAlaGlyLysThrGlyProGlyGlyProProGlyProValGlyProAlaGly 233
QY 168 TCACAGCTGCTGGCGCCGACCGGGA-----AGTTGGCGATCTCTCCATCTTCCCGC 115
Db 234 ProSerGlyLysProGlyAlaProGlyGlnProGlyProHisGlyProProGlyGlnPro 253
QY 114 GCAGAAAGCGCGGACCGCCAGCCATGGAACCTGGAAGTACTGCTCTCGA 64
Db 254 GlyGlnAspAlaGlnTyrCysProCysProArgSerLeuCysSerArg 270

RESULT 94
A:1219
collagen 1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 30-Jun-1989 #sequence revision 30-Jun-1989 #text_change 09-Jul-2004
C:Accession: A31219; T37290; T37291
R:Kramer, J.M.; Cox, G.N.; Hirsh, D.
Cell 30, 599-606, 1982
A:Title: Comparisons of the complete sequences of two collagen genes from Caenorhabditis
A:Reference number: A30826; MUID:83050944; PMID:7139711
A:Accession: A31219
A:Molecule type: DNA
A:Residues: 1-296 <KRA>
A:Cross-references: UNIPROT:P08124; UNIPARC:UPI0000127203; GB:V00147; NID:G6677; PIDN:CA
A:Accession: T37290
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-296 <KR2>
A:Cross-references: UNIPARC:UPI0000127203; EMBL:J01047; PIDN:AAA27988.1
A:Accession: T37291
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-296 <KR3>
A:Cross-references: UNIPARC:UPI0000127203; EMBL:V00147; NID:G6677; PIDN:CAA23463.1; PID:
C:Genetics:
```

```
A:Gene: col-1
A:Introns: 54/3; 280/2

Alignment Scores:
Pred. No.: 0.00815 Length: 296
Score: 144.50 Matches: 68
Percent Similarity: 28.4% Conservative: 10
Best Local Similarity: 24.7% Mismatches: 92
Query Match: 3.3% Indels: 105
DB: 2 Gaps: 13

US-10-768-158-1 (1-2419) x A31219 (1-296)
QY 780 TCCACAGCCCAACTCTTCCCGCCACGCGGCGCTCAGCGT----- 736
Db 75 ThrThrArgGlnAlaTyrGlyGlyProGluValAsnProAlaProAsnLeuGlnCysGlu 94
QY 735 -----TCGACACTGGTCCACCACTGGTGGCAGTGTCTCCGTCA----- 697
Db 95 GlyCysCysLeuProGlyProProGlyProAlaGlyAlaProGlyLysProGlyLysPro 114
QY 696 GGGCTTCCAGCTGGGCTTGTACAGGACACCCCGGAGGAATCTGGCCAGCTGCTCCA--- 640
Db 115 GlyArgProGlyAlaProGlyThrProGlyLysProProValAlaProCys 134
QY 639 ---CCATCGTCACCAAGTCCCGATGTCATCTTATCTTATCTTGAAGGAGTCTCGGTAGTCTAGTCCGAGAG 583
Db 135 GluProThrThrProProProCysLysProCysProGlnGlyProProGlyProProGly 154
QY 582 CCATCGGTGTCTCCAGAACTCTTCACAGTGTCTCAACACGAGGAGCGGTAGCCAGCTTAT 523
Db 155 ProProGlyAlaProGlyAspProGlyGluAlaGlyThrProGlyArg----- 170
QY 522 CATTCAAACTCCGCGAGAATTCTTGAAGGTGCTCGGTAGTCTAGTCTAGTCCGAGAG 463
Db 171 -----ProGlyThrAspAlaAla 176
QY 462 AGCGGTGAAGTATATAAGACACACCAGATCTTGGGGTTGGAGCCCATATAGATGA 403
Db 177 ProGlySerProGlyProArgGlyProProGlyPro----- 188
QY 402 CTTGGAGTCTCCATTGTGGAGGTGAGAGGCGAGAAAGCGGTAGGCGAGGTGGCTCTTGA 343
Db 189 -----AlaGlyGluAlaGly----- 193
QY 342 TGAGCGGGGAGAGTCAAGTTCCTTGTATGATGTCAGGCCCGGCTGTGGGTACTTCCAGA 283
Db 194 -----AlaProGlyProAla---GlyGluProGly 202
QY 282 CCGGAGCTGCTGCTGATGTTTCATCAAGCCGATCTCATCGGGTTCAGGCGCTGCTCA 223
Db 203 ThrProAlaIleSerGluProLeuThr-----ProGlyAla 214
QY 222 CCAAGTAGACCA-----CCTCTCGCAGCAAGCTGGTGGCGGACT 184
Db 215 ProGlyGluProGlyAspSerGlyProProGlyProProGlyProPro----- 230
QY 183 TGGGTAGGTGACGATCCACACAGTCTGGTGGCCCGCCAGCGGAAGTTGGCGATCTCTCCA 124
Db 231 -----GlyAlaProGlyAsn---AspGlyProPro 239
QY 123 TCTTCCCGGCGAGAAAGCGGCGAGCCGACCGCATCGAACTCGAAGTACTTGCTCTCGA 64
Db 240 GlyProProGlyProLysGlyAlaProGlyProAspGly----- 252
QY 63 ACTCCCGCGGGTGTGGGGG---TCTCGGCTCTGCTCTCGGCCA 22
Db 253 ---ProProGlyAlaAspGlyGlnSerGlyProProGlyProPro 266

RESULT 95
CGHU6C
N:Alternate names: procollagen alpha 1(II) chain
collagen alpha 1(II) chain precursor [validated] - human
```

N;Contains: chondrocalcin; collagen alpha 1(II) chain precursor splice form 1; collagen
 C:Species: Homo sapiens (man)
 C:Date: 28-May-1986 #sequence revision 01-Sep-1995 #text change 31-Dec-2004
 A:Cross-references: A38513; S06715; S24270; A24828; S06496; A35428; A30147; A33116; S64674; S63750; I37251; I37252; I37253; I37254; I55338; I59535; I61910
 R:Ryan, M.C.; Sieraski, M.; Sandell, L.J.
 Genomics 8, 41-48, 1990
 A:Title: The human type II procollagen gene: identification of an additional protein-cod
 A:Reference number: A38513; MUID:91184811; PMID:2081599
 A:Accession: A38513
 A:Molecule type: DNA
 A:Residues: 1-103 <RYA>
 A:Cross-references: UNIPROT:P02458; UNIPROT:Q14042; UNIPROT:Q16672; UNIPROT:Q12985; UNIP
 R:Su, M.W.; Lee, B.; Ramirez, F.; Machado, M.; Horton, W.
 Nucleic Acids Res. 17, 9473, 1989
 A:Title: Nucleotide sequence of the full length cDNA encoding for human type II procolla
 A:Reference number: S06715; MUID:90067946; PMID:2587267
 A:Accession: S06715
 A:Molecule type: mRNA
 A:Residues: 1-28, 'R', '99-1487 <SU2>
 A:Cross-references: UNIPARC:UPI0000173B64; EMBL:X16468; NID:g29515; PIDN:CAA34488.1; PID
 A:Note: alternative splice form 1
 R:Vikula, M.; Metsaeranta, M.; Syvaenen, A.C.; Ala-Kokko, L.; Vuorio, E.; Peltonen, L.
 Biochem. J. 285, 287-294, 1992
 A:Title: Structural analysis of the regulatory elements of the type-II procollagen gene.
 A:Reference number: S24270; MUID:92344585; PMID:1637314
 A:Accession: S24270
 A>Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-28 <VIK>
 A:Cross-references: UNIPARC:UPI0000173B64; EMBL:X58709; GB:S40537; NID:g35659
 A:Note: this translation is not annotated in Genbank entry HSPROCOB1, release 111.0
 R:Numaz, A.M.; Kohno, K.; Martin, G.R.; Yamada, Y.
 Gene 44, 11-16, 1986
 A:Title: Promoter region of the human pro-alpha-1-(II)-collagen gene.
 A:Reference number: A24828; MUID:87031574; PMID:30211582
 A:Accession: A24828
 A:Molecule type: DNA
 A:Residues: 1-8, 'T', '10-28 <NUN>
 A:Cross-references: UNIPARC:UPI000016A71A; GB:M25698; NID:g180872; PIDN:AAAS2051.1; PID:
 R:Baldwin, C.T.; Reginato, A.M.; Smith, C.; Jimenez, S.A.; Prockop, D.J.
 Biochem. J. 262, 521-528, 1989
 A:Title: Structure of cDNA clones coding for human type II procollagen. The alpha-1(II)
 A:Reference number: S06496; MUID:90026318; PMID:2803268
 A:Accession: S06496
 A:Molecule type: mRNA
 A:Residues: 7-28, 'R', '99-157, 'P', '159-440, 'G', '442-456, 'E', '458-640, 'A', '642-831, 'PA', '834, 'F'
 A:Cross-references: UNIPARC:UPI0000173B65; EMBL:X16711; NID:g30040; PIDN:CAA34683.1; PID
 A:Note: alternative splice form 1
 R:Ryan, M.C.; Sandell, L.J.
 J. Biol. Chem. 265, 10334-10339, 1990
 A:Title: Differential expression of a cysteine-rich domain in the amino-terminal propept
 A:Reference number: A35428; MUID:90285153; PMID:2355003
 A:Accession: A35428
 A>Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 27-81, 'L', '83-103 <RYA2>
 A:Cross-references: UNIPARC:UPI0000173B66
 A:Note: alternative splice form 2; splicing appears to be under developmental regulation
 R:Su, M.W.; Benson-Chanda, V.; Vissing, H.; Ramirez, F.
 Genomics 4, 438-441, 1989
 A:Title: Organization of the exons coding for Pro alpha-1(II) collagen N-propeptide conf
 A:Reference number: A30147; MUID:89233138; PMID:2714801
 A:Accession: A30147
 A:Molecule type: DNA
 A:Residues: 104-157, 'P', '159-236 <SUM>
 A:Cross-references: UNIPARC:UPI0000173B67; GB:J03065; GB:M23660; GB:M25655; GB:M25656; G
 R:Ala-Kokko, L.; Baldwin, C.T.; Moskowitz, R.W.; Prockop, D.J.
 Proc. Natl. Acad. Sci. U.S.A. 87, 6565-6568, 1990
 A:Title: Single base mutation in the type II procollagen gene (COL2A1) as a cause of pri
 A:Reference number: A94227; MUID:90370826; PMID:1975693
 A:Accession: A33116

A:Molecule type: DNA
 A:Residues: 171-172, 'C', '174-175 <ALA>
 A:Cross-references: UNIPARC:UPI0000173B68
 A:Note: mutant sequence from a family with family with primary generalized osteoarthritis;
 R:Diab, M.; Wu, J.J.; Eyre, D.R.
 Biochem. J. 314, 327-332, 1996
 A:Title: Collagen type IX from human cartilage: a structural profile of intermolecular
 A:Reference number: S64673; MUID:96195147; PMID:8660302
 A:Accession: S64674
 A:Molecule type: protein
 A:Residues: 188-189, 'X', '191-195,1224-1230, 'X', '1232-1236 <DIA>
 A:Cross-references: UNIPARC:UPI0000173B69; UNIPARC:UPI0000173B6A
 R:Franc, S.; Marzin, E.; Boutillon, M.M.; Lafont, R.; Lechene de la Porte, P.; Herbage,
 Eur. J. Biochem. 234, 125-131, 1995
 A:Title: Immunohistochemical and biochemical analyses of 20000-25000-year-old fossil cal
 A:Reference number: S63514; MUID:96096730; PMID:8529631
 A:Accession: S63514
 A:Molecule type: protein
 A:Residues: 243-261,575-590,756-763, 'X', '765-779 <FRA>
 A:Cross-references: UNIPARC:UPI0000173B6B; UNIPARC:UPI0000173B6C; UNIPARC:UPI0000173B6D
 R:Tiller, G.E.; Weis, M.A.; Polunbo, P.A.; Gruber, H.E.; Rimoin, D.L.; Cohn, D.H.; Eyre,
 Am. J. Hum. Genet. 56, 388-395, 1995
 A:Title: An RNA-splicing mutation (G+5IVS20) in the type II collagen gene (COL2A1) in a
 A:Reference number: I38867; MUID:95150028; PMID:7847372
 A:Accession: I38867
 A>Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: DNA
 A:Residues: 440, 'G', '442-456, 'E', '458-480, 'P', '482-509 <TIL1>
 A:Cross-references: UNIPARC:UPI000006F3AF; EMBL:U15195; NID:g557053; PIDN:AAB60370.1; PI
 R:Ramirez, F.
 submitted to the EMBL Data Library, December 1988
 A:Reference number: S04892
 A:Accession: S04892
 A:Molecule type: mRNA
 A:Residues: 501-676, 'A', '678-783, 'A', '785-831, 'PA', '834, 'F', '836-1214 <RAM>
 A:Cross-references: UNIPARC:UPI000016A700; EMBL:X13783; NID:g30037; PIDN:CAA32030.1; PI
 R:Vikula, M.; Peltonen, L.
 FEBS Lett. 250, 171-174, 1989
 A:Title: Structural analyses of the polymorphic area in type II collagen gene.
 A:Reference number: S05000; MUID:89325561; PMID:2753125
 A:Accession: S05000
 A:Molecule type: DNA
 A:Residues: 630-640, 'A', '642-785 <VIK2>
 A:Cross-references: UNIPARC:UPI0000173B6E; EMBL:X16158; NID:g29951; PIDN:CAA34278.1; PID
 A34282.1; PID:g1335022; PIDN:CAA34283.1; PID:g1335023; PIDN:CAA34284.1; PID:g1335024
 R:Bogaert, R.; Tiller, G.E.; Weis, M.A.; Gruber, H.E.; Rimoin, D.L.; Cohn, D.H.; Eyre, D.
 J. Biol. Chem. 267, 22522-22526, 1992
 A:Title: An amino acid substitution (Gly853-->Glu) in the collagen alpha 1(II) chain pro
 A:Reference number: A44309; MUID:93054548; PMID:1429602
 A:Accession: A44309
 A>Status: nucleic acid sequence not shown; not compared with conceptual translation
 A:Molecule type: DNA; mRNA
 A:Residues: 752-831, 'PA', '834, 'F', '836-1005, 'K', '1007-1036, 'Q', '1038-1052, 'E', '1054-1068, 'T',
 A:Cross-references: UNIPARC:UPI0000173B6F; GB:L00977; NID:g180812; PIDN:AAB23914.1; PID:
 A:Note: sequence extracted from NCBI backbone (NCBI:117273); parts of this sequence wer
 A:Note: this translation is not annotated and this publication is not cited in GenBank e
 A:Note: mutant sequence associated with perinatal lethal hypochondrogenesis
 R:Tiller, G.E.; Rimoin, D.L.; Murray, L.W.; Cohn, D.H.
 Proc. Natl. Acad. Sci. U.S.A. 87, 3889-3893, 1990
 A:Title: Tandem duplication within a type II collagen gene (COL2A1) exon in an individua
 A:Reference number: S16502; MUID:90251662; PMID:2339128
 A:Accession: S16502
 A:Molecule type: DNA
 A:Residues: 1164-1184, 'GPSKDGANGIPQPI', '1185-1199 <TIL2>
 A:Cross-references: UNIPARC:UPI000011F7F2; EMBL:M37126; NID:g180808; PIDN:AAA52037.1; PI
 A:Note: mutant sequence from a patient with spondyloepiphyseal dysplasia
 R:Cheah, K.S.E.; Stoker, N.G.; Griffin, J.R.; Grosveld, F.G.; Solomon, E.
 Proc. Natl. Acad. Sci. U.S.A. 82, 2555-2559, 1985
 A:Title: Identification and characterization of the human type II collagen gene (COL2A1)
 A:Reference number: A02858; MUID:85190534; PMID:3857598
 A:Accession: A02858
 A:Molecule type: DNA
 A:Residues: 1032-1056, 'N', '1058-1068, 'T', '1070-1487 <CHE>

A;Cross-references: UNIPARC:UPI000016A6BB; GB:J00116; NID:g180395; PIDN:AAA51997.1; PID:R;Elima, K.; Vuorio, T.; Vuorio, E.
Nucleic Acids Res. 15, 9499-9504, 1987
A;Title: Determination of the single polyadenylation site of the human pro-alpha-1(II) o
A;Reference number: A27280; MUID:8806771; PMID:2825137
A;Accession: A27280
A;Molecule type: DNA; mRNA
A;Residues: 1175-1487 <BLI>
A;Cross-references: UNIPARC:UPI000016A71B; EMBL:X06268; NID:g30096; PIDN:CAA29604.1; PID:
A;Experimental source: fetal epiphyseal cartilage
R;van der Rest, M.; Rosenberg, L.C.; Olsen, B.R.; Poole, A.R.
Biochem. J. 237, 923-925, 1986
A;Title: Chondrocalcin is identical with the C-propeptide of type II procollagen.
A;Reference number: A57033; MUID:87099927; PMID:3800925
A;Accession: A57033
A;Molecule type: protein
A;Residues: 'XE',1244-1246,'N',1248,'X',1250-1265,1295-1305,1395-1408 <VAN>
A;Cross-references: UNIPARC:UPI0000173B70; UNIPARC:UPI0000173B71; UNIPARC:UPI0000173B72
A;Note: chondrocalcin identified as released collagen 1(II) chain carboxyl-terminal prope
R;Strom, C.M.; Upholt, W.B.
Nucleic Acids Res. 12, 1025-1038, 1984
A;Title: Isolation and characterization of genomic clones corresponding to the human TYP
A;Reference number: A21733; MUID:84118798; PMID:6320112
A;Accession: A21733
A;Molecule type: DNA
A;Residues: 1245-1295 <STR1>
A;Cross-references: UNIPARC:UPI000016A61A; EMBL:X00339; EMBL:X00298; NID:g394699; PIDN:C
A;Accession: 821733
A;Molecule type: DNA
A;Residues: 894-909,'PE' <STR2>
A;Cross-references: UNIPARC:UPI000006EA40; GB:K01785; NID:g30035; PIDN:CAA25082.1; PID:G
R;Nunez, A.M.; Francomano, C.; Young, M.F.; Martin, G.R.; Yamada, Y.
Biochemistry 24, 6343-6348, 1985
A;Title: Isolation and partial characterization of genomic clones coding for a human pro
gene.
A;Reference number: A24561; MUID:86104139; PMID:3002437
A;Accession: A24561
A;Molecule type: DNA
A;Residues: 1296-1358 <NUN2>
A;Cross-references: UNIPARC:UPI0000173B73; GB:M12048; NID:g180017
A;Note: this translation is not annotated in GenBank entry HUMCCT2A, release 111.0
A;Note: the codons given for 1333-Lys (AAG) and 1350-Gly (GCA) are inconsistent with the
R;Sangioirgi, F.O.; Benson-Charanda, V.; de Wet, W.J.; Sobel, M.E.; Tsipouras, P.; Ramirez,
Nucleic Acids Res. 13, 2207-2225, 1985
A;Title: Isolation and partial characterization of the entire human pro alpha 1(II) coll
A;Reference number: I3749; MUID:85215609; PMID:2987845
A;Accession: S59491
A;Molecule type: DNA
A;Residues: 7-28,'R',99-114;541-578;786-802;1055-1056,'N',1058-1068,'T',1070-1109;1200-1
A;Cross-references: UNIPARC:UPI0000071780; UNIPARC:UPI0000173B74; UNIPARC:UPI0000173B75;
A;Accession: 184453
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 7-28 <SAN2>
A;Cross-references: UNIPARC:UPI0000071780; UNIPARC:UPI0000173B74; UNIPARC:UPI0000173B75;
24938; NID:g30104
A;Note: the GenBank PID is based on an incorrect reading frame

Alignment Scores:
Pred. No.: 0.0126 Length: 1487
Score: 144.00 Matches: 78
Percent Similarity: 33.2% Conservative: 20
Best Local Similarity: 26.4% Mismatches: 131
Query Match: 3.3% Indels: 66
DB: 1 Gaps: 15

US-10-768-158-1 (1-2419) x CGHU6C (1-1487)

Qy 759 GGCCACGGGAGGGCCCTCAGCGTTGCAGCACTGTCACACGCTGGTGGCAGTGCTCCG 700
|||||
Db 207 GlyProArgGlyProPro-----GlyProAlaGlyAlaProGlyPro 220
699 TCAGGGCTTCACGCTGGGCGCTTGTCACAGGACACCCAGGAATC----- 655

A:Introns: 54/3; 285/2

Alignment Scores:

Pred. No.: 0.00974 Length: 301
Score: 143.50 Matches: 68
Percent Similarity: 28.4% Conservative: 10
Best Local Similarity: 24.7% Mismatches: 92
Query Match: 3.3% Indels: 105
DB: 2 Gaps: 13

US-10-768-158-1 (1-2419) x T21314 (1-301)

```
QY 780 TCCACAGCCCACTCTTCCCGGCCACCGGCGAGGCGCTCAGCGT----- 736
Db 80 ThrThrArgGlnAlaTyGlyGlyProGluValAsnProAlaProAlaAsnLeuGlnCysGlu 99
QY 735 -----TGCAAGCACTGGTCCACAGCTGGTGGCAGTCTCCGTCA----- 697
Db 100 GlyCysCysLeuProGlyProGlyProAlaGlyAlaProGlyLysProGlyLysPro 119
QY 696 GGGCTTCCAGCTGGGCGTGTGCACAGGACACCCCGAGGAATCTGGCCAGCTGCTCCA--- 640
Db 120 GlyArgProGlyAlaProGlyThrProGlyThrProGlyLysProProValAlaProCys 139
QY 639 ---CCATCGTCACCGTCCCATGCTCTTCATCTAGACAAAACACGCTTCGAGT 583
Db 140 GluProThrThrProProProCysLysProCysProGlnGlyProProGlyProGly 159
QY 582 CCATCGGTGCTCCCAAGAACTCTGCACGTGTCAAAACGAGGCGGTAGCCAGCTTAT 523
Db 160 ProProGlyAlaProGlyAspProGlyGluAlaGlyThrProGlyArg----- 175
QY 522 CATTCATAAACCTCCGGCAGAAATCTTGAAGGTGCTCGTAGCTCATGTGCCAGAG 463
Db 176 -----ProGlyThrAspAlaAla 181
QY 462 AGCGGTGGAAGTATAATAAGACACACACAGATCTCTTGGGTGTCAGGACCATATAGTA 403
Db 182 ProGlySerProGlyProArgGlyProProGlyPro----- 193
QY 402 CCTTGGAGTCTCATTTGTGAGGTGAGAGGCGAGAAAGCGGTAGGCGAGTGTCTTGA 343
Db 194 -----AlaGlyGluAlaGly----- 198
QY 342 TGAGGCGGGAGAGTCACTTCTTGATGATGTCAGGCGCGGCTGTGGTACTCCAGGA 283
Db 199 -----AlaProGlyProAla---GlyGluProGly 207
QY 282 CCGGAGTGTGCTGTCGATGTTTCATCAAGCCGATCTATCGGGGTGAGCGCCCTGCTCA 223
Db 208 ThrProAlaIleSerGluProLeuThr-----ProGlyAla 219
QY 222 CCAAGTAGACCA-----CCTCTGCAGCAAGCTGGTCCGGACT 184
Db 220 ProGlyGluProGlyAspSerGlyProProGlyProProPro----- 235
QY 183 TGGGGTAGGTGACGATCCACACGCTCGCTGGGCGGACCGGAAAGTTGGCGATCTCTCCA 124
Db 236 -----GlyAlaProGlyAsn---AspGlyProPro 244
QY 123 TCTTCCCGGCGAGAAAGCGGCGAGCCGACGCGCATATGGAATCTGCTCTCGCA 64
Db 245 GlyProProGlyProLysGlyAlaProGlyProAspGly----- 257
QY 63 ACTCCCGCGGGTGTGGGG---TCTCGGCTCGCTCTCGCCA 22
Db 258 ---ProProGlyValAspGlyGlnSerGlyProProGlyProPro 271
```

RESULT 97

T20497

hypothetical protein F02D10.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T20497

R:Swinsburne, J.

submitted to the EMBL Data Library, November 1995

A:Reference number: Z19283

A:Accession: T20497

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-316 <WIL>

A:Cross-references: UNIPROT:Q19111; UNIPARC:UPI0000080924; EMBL:Z67990; PIDN:CAA91932.1;

A:Experimental source: clone F02D10

C:Genetics:

A:Gene: CESP:F02D10.1

A:Map position: X

A:Introns: 56/3

Alignment Scores:

Pred. No.: 0.00984 Length: 316
Score: 143.50 Matches: 66
Percent Similarity: 30.1% Conservative: 8
Best Local Similarity: 26.8% Mismatches: 87
Query Match: 3.3% Indels: 85
DB: 2 Gaps: 13

US-10-768-158-1 (1-2419) x T20497 (1-316)

```
QY 735 TGCAGCACTGGTCCACAGCTGGTGGCAGTGTCTCCGTCA----- 697
Db 117 CysAsnProGlyProProGlyProGlyArgProGlyLysProGlyThrProGlyLys 136
QY 696 ---GGCTTCCA-----GCTGGGCGCTGTCTCAC--- 673
Db 137 ProGlyAlaProGlyAsnProGlyAlaSerGlyLysGlyAlaAlaProCysGluPro 156
QY 672 AGGACACCCCGAGGAATC-----TGGCCAGCTGTCCACCATCTGCACAGTCCCGAT 619
Db 157 SerThrProProProCysGlnProCysProAlaGlyProProGlyProProGlyPro 176
QY 618 GCATGCTTTCATCTTGAGAAAACACGTTTCGAGTCCATCGGTGCTCCAGAGACTCT 559
Db 177 GlyThr-----ProGlyGluProGlyProGlyProAlaGlySerProAlaGlyPro 194
QY 558 GCACGTGTCAAAACAGGAGCGGTAGCCAGCTTATCATTAACCTCCGCGCAGAAAT 499
Db 195 Ser----- 195
QY 498 CTTGAAAGGTGCTCGTAGCTCATGTCCGAGAGAGCGGTGGAACATGATAATAAGACA 439
Db 196 -----GlyProGlyProAlaGly 201
QY 438 CCACAGATCTTGGGTTGCGAGCCATATAGATGACCTTGAGTCTCCATTTGGAGGT 379
Db 202 ProProGlyProAlaGly-----Pro-----ProGlyAsnAspGlyGlnProGly 216
QY 378 CAGAGGCGAGAAAGCGGTAGGCGAGTGGCTTTCATGAGCGGGGAGAGGTGAGTCTCT 319
Db 217 GlnProGlyGlyProGlyGlnAspGlyAlaSerSerAlaGlyGlyGluAlaGly----- 234
QY 318 TGATGATGTCAGGCGCGGCTGTGGTACTCCAGGACCG---GGAGCTGCTCGTCGATGT 262
Db 235 -----ProGlyProAla---GlyProProGlyProAlaGlyProAlaGlyPro 250
QY 261 TCATCAAGCCGATCTCATCGGGGTGAGCGCTTGCTCACCAGTAGACCTCTCTGCA 202
Db 251 GlyGlnSerGlySerGlySerAlaGlyGlyProGlyProLysGlyProProGlyProAla 270
QY 201 GCAAGCTGTGCGGAGTGTGGGTAGGTGACGATCCACACGCTGCTGGCGCGCAGCGGA 142
Db 271 GlyGln-----ProGly 274
QY 141 AGTTGGCGATCTCTTCATCTTCCGCGGCGAGAGGGCGGCGGACCGCATGGAAC 82
Db 275 Ser---AspGlyAsnProGlyThrAlaGlyProProGlyAsnProGlyGlyGluGly 293
```


Score:	143.50	Matches:	82
Percent Similarity:	31.6%	Conservative:	11
Best Local Similarity:	27.9%	Mismatches:	102
Query Match:	3.3%	Indels:	99
DB:	1	Gaps:	18
US-10-768-158-1 (1-2419) x CGRTIS (1-671)			
Qy	762	CCCGGCCACCGGCGAGGGCCTCAGCGTTGCAGCACTGGTCCACCAAGCT	----- 715
Db	289	ProGlyPro-----AlaGlyValGln-----GlyProProGlyProAlaGlyGlu	303
Qy	714	-----GGTGGCAGTGCTCGGTCCAGGGCTCCAGCTGGGCGCTTGT	676
Db	304	GluglyLysArgGlyAlaArgGlyGluProGlyProSerGlyLeuProGlyProGly	323
Qy	675	CACAGGACACCCCGAGGAATC-----	655
Db	324	GluArgGlyGlyProGlySerArgGlyPheProGlyAlaAspGlyValAlaGlyProLys	343
Qy	654	TGGCCAGCTGCTCCACCATCGTCACCAAGTCCCGATGCATGCTTTCATACTTGAGAAAAA	595
Db	344	GlyProAlaGlyGluArgGlySerProGlyPro-----	354
Qy	594	GCAGTTGAGTCCATCGGTGCTCCAGAACTCTGCAGTGCTCAACACGAGGCCGT	535
Db	355	-----AlaGlyProLysGlySerProGlyGluAlaGlyArgProGlyGluAlaGly	371
Qy	534	AGCCAGCTTATCATTCATAAACCTCCGGCAGAAATCTTGAAAGGTGCCTCGTAGCTCA	475
Db	372	-----LeuProGlyAlaLysGlyLeuThr	379
Qy	474	TGTTCCGACAGAGCGGTGGAACATGATAATAAGACA-----CCACCAAGATCCTTTGGGGTTGC	418
Db	380	GlySerProGlySerProGlyProAspGlyLysThrGlyProProGlyProAlaGly	398
Qy	417	GAGCCATATAGATCACTTGAGAGTCTCCATTGTGGAGTCCAGGGCAGAAAGC-----	364
Db	399	*****GlyArg-----ProGlyProAlaGlyProProGlyAlaArgGlyGlnAlaGlyVal	417
Qy	363	-----GGTAGGCGAGTGGCTCTTGATGAGGGGGGAGAGGTCA	325
Db	418	MetGlyPheProGlyProLysGlyThrAlaGlyGluProGlyLysAlaGlyGluArgGly	437
Qy	324	GTTCCTTGATGATCTCCAGGCCCG-----	301
Db	438	Val-----ProGlyProProGlyAlaValGlyProAlaGlyLysAspGlyGlu	453
Qy	300	-----GCTGTGGTACTCCAGGACCGGAGCTGCTCGTAGTTTCATCAAGCCGATCT	247
Db	454	AlaGlyAlaGlnGlyAlaProGlyProAlaGlyPro-----AlaGlyGluArgGly	470
Qy	246	CATCGGGTTCAG-----CGCCCTGGCTCCACCAAGTAGACCACTCTGTCGCAAGCTGG	193
Db	471	GluGlnGlyProAlaGlySerProGlyPheGlnGlyLeuProGlyProAlaGly	488
Qy	192	TGCCGAGCTTGGGTAGGTGACGATCCACAGTCCGCTGGGCGGCACAGGAAAGTTGGCGA	133
Db	489	-----ProProGlyGluAlaGlyLysProGly*****GlyValProGlyAspLeuGly	506
Qy	132	TCTCTCT---CCATCT---TCCCGCGGCAGAGGGCGGCGGACCGCATGGAAGTCA	79
Db	507	AlaProGlyProSerGlyAlaArgGlyGluArgGlyPheProGlyGluArgGlyValGln	526
Qy	78	AGTACTTGCTCTCGAACTCCCGCGGGTGTGGGGTCTCGG	37
Db	527	Gly-----ProProGlyProAlaGlyProArg	535

RESULT 100

JQ0405

RESULT 100
JQ0405
hypothetical 119.5K protein (uvrA region) - Micrococcus luteus
N;Alternate names: ORF 1 protein
C:Species: Micrococcus luteus. Micrococcus lysodeikticus

C>Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 04-Feb-2000
C:Accession: JQ0405
R:Shiota, S.; Nakayama, H.
Mol. Gen. Genet. 217, 332-340, 1989
A>Title: Micrococcus luteus homolog of the Escherichia coli uvra gene: identification of
A:Reference number: S04781; MUID:89364717; PMID:2549377
A:Accession: JQ0405
A:Molecule type: DNA
A:Residues: 1-1106 <SHI>
A:Cross-references: UNIPARC:UPI0000177394; EMBL:X15867
A>Note: this reading frame extends between two stop codons and does not begin with a sta
A>Note: the gene encoding this protein overlaps uvra gene
C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;

Alignment Scores:
Pred. No.: 0.0129 Length: 1106
Score: 143.50 Matches: 102
Percent Similarity: 34.2% Conservative: 17
Best Local Similarity: 29.3% Mismatches: 91
Query Match: 3.3% Indels: 139
DB: 2 Gaps: 22

US-10-768-158-1 (1-2419) x JQ0405 (1-1106)

```
QY 5 CGCGCAGCGCGCGCATGCGCGAGAGCGAGCGCGAGACCCCGACACCCCGGGGAGTT 64
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
317 ArgValArgGlyArgGlyAlaGlyArgGluGlyGlnGluGluHis---ArgGlyVal 335
QY 65 CGAGAGCAAGTACTTCGA-----GTTTCATGCGCGTGCCTGCCCGCTTCTG 112
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
336 ArgArgAlaGlyArgGlyGlnProValProLeuValLeuArgGluAlaLeuLeu 355
QY 113 CGCGGGAGAGTAGGAGAGATCGCAACTTCCCGTGGCGCC-----CAGCGA 160
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
356 ProGlnArgAlaArgAlaAspArgGlyArgAspArgAlaAlaLeuValLeuGlnGln 375
QY 161 CQTGTGGATCGACACTACCCCAAGTC---CGSCACCGACTTGTGCGAGGAGTGTCTA 217
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
376 ProVal---ArgArgValProArgValHisArgHisArgLeuProProAlaGlyGly--- 393
QY 218 CTTGTGTAGCGCGCGTGCACCGCATGAGATCGGCTTGATGAACATCGACGACGCT 277
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
394 -----ProGly-----ProArg-----ArgArgGlnArgAlaVal 404
QY 278 CCGGTT-----283
Db |||||
405 ProAlaArgGlyArgArgAlaValValAlaArgGlnValHisLeuGlyLeuAla 424
QY 284 -----CCTGGAGTACCCACACGCGGG 304
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
425 AlaArgAlaArgAlaGlyGlnGlyAspGlyLeuProGlyHisProValGluGly 444
QY 305 CCT-----GGACATCATCAAGAACTGAC 328
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
445 ProAspGlyGlyAlaArgArgProAlaArgGlnGlyLeuGlnGlyGlyAsp 464
QY 329 CTCTCCCGCCTCATCAAGAGCCACTGCCCTACCGCTTCTGCCCTCTGACCTCCAA 388
Db ::|||::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
465 Val-ProGlnProValArgProAlaPro-----LeuHis 477
QY 389 TGGAGACTCCAAGTCACTATATGGCTCGCAACCCCAAGGATCTGGTGGTCTTATTA 448
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
477 eGly-----LeuArgGlyArgHisPro-----484
QY 449 TCAGTTCCACCGCTCTTCGGACCATGAGCTACCGAGCACCTTTCAAGAATTTCGCG 508
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
485 -----LeuArgAspAlaGlnAlaArgGly-----492
QY 509 GAGGTTTATGATGATACTGGGCTA-----CGGCTCTGGTTTGAGCACGTGCAGG 561
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
493 -----AspArgValGlyArgArgProArgProAlaLeuArgValValHisAlaG 508
QY 562 A-----GTTCTGGGAGCA-----CGCATGGACTCGAACGTGCTTT 597
```

```
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
508 YAspProValProGlyValProArgGlyProProGlnProHisGlyProGluArgAlaAr 528
QY TTCTCAA---GTATGAAGA-----CATGCATCGGACCTGGT-----631
Db ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
528 gGlyArgProValHisArgGlyArgHisProProAlaHisAlaArgGlyHisGlyVal 548
QY 632 -----GACGATGTTGGAGCAGCTGCCAGATTCTCTGGGGTGTCTCTGTG 675
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
548 uLeuGlyAlaAlaAlaAspGlyProGlyAlaAlaAspArgGlyProGlyAlaGlnGlyAs 568
QY 676 ACAAGGCCACGCTGGAAGCCCTGACGGAGCACTGCCA-----CCAGC 717
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
568 pProGlyProAlaGlyValProAlaGlyArgArgProArgValProGlnProArgAlaAl 588
QY 718 TGGTGGACCACTGTCGAACCGCTGAGCCCTGCCGCGCGCGGGAAGAGATTGGGCTGT 777
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
588 aGlyArgHisProLeuArgArgArgGlyProAlaHisProPro-----602
QY 778 GGAAGGACATCTTCACCGTCTC 799
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
603 ----GlyHisThrAspArgLeu 608
```

Search completed: May 17, 2006, 10:00:17
Job time : 252.5 secs


```

Sequence 31, Appl
Sequence 32, Appl
Sequence 87, Appl
Sequence 13, Appl
Sequence 518, App
Sequence 97, Appl
Sequence 40, Appl
Sequence 393, App
Sequence 97, Appl
Sequence 2, Appl
Sequence 2466, Ap
Sequence 7, Appl
Sequence 17, Appl
Sequence 54, Appl
Sequence 2959, Ap
Sequence 466, App
Sequence 21, Appl
Sequence 224, App

83      83      1.9      558      7      US-11-311-754-31
84      83      1.9      811      6      US-10-505-928-32
85      83      1.9      811      6      US-10-505-928-87
86      83      1.9      951      6      US-10-199-229-13
87      83      1.9      1137     6      US-10-196-749-518
88      83      1.9      1894     6      US-10-196-749-97
89      82.5     1.9      665      6      US-10-370-959-40
90      82.5     1.9      667      6      US-10-505-928-393
91      82.5     1.9      1894     6      US-10-196-749-97
92      82.5     1.9      4373     7      US-11-118-524-2
93      82      1.9      362      6      US-10-511-937-2466
94      82      1.9      872      6      US-10-519-335-7
95      82      1.9      872      7      US-11-302-678-17
96      81.5     1.9      384      6      US-10-505-928-54
97      81.5     1.9      620      6      US-10-511-937-2959
98      81.5     1.9      813      6      US-10-196-749-466
99      81.5     1.9      1763     6      US-10-504-120-21
100     81      1.8      341      6      US-10-196-749-224

ALIGNMENTS

RESULT 1
US-10-985-570-2
; Sequence 2, Application US/10985570
; Publication No. US20060100138A1
; GENERAL INFORMATION:
; APPLICANT: Olsen, David R.
; APPLICANT: Polarek, James W.
; APPLICANT: Yang, Chunlin
; TITLE OF INVENTION: IMPLANTABLE COLLAGEN COMPOSITIONS
; FILE REFERENCE: FP0305 US
; CURRENT APPLICATION NUMBER: US/10/985,570
; CURRENT FILING DATE: 2004-11-10
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 1464
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-985-570-2

Alignment Scores:
Pred. No.: 9,136-06 Length: 1464
Score: 155.50 Matches: 86
Percent Similarity: 29.8% Conservative: 11
Best Local Similarity: 26.5% Mismatches: 107
Query Match: 3.5% Indels: 121
DB: 6 Gaps: 20

US-10-768-158-1 (1-2419) x US-10-985-570-2 (1-1464)

QY 783 CTTTCCACAGCCAACTCTTCCCGGCCCA-----CGGCAGGCGCTCAGCGTTGCAGC 730
Db 870 ProProGlyAlaThrGlyPheProGlyAlaAlaGlyArgValGlyProProGlyProSer 889
QY 729 -----ACTGTCACACAGCT-----GGTGGCAGTGCCTCCG 700
Db 890 GlyAenAlaGlyProProGlyProProGlyProAlaGlyLysGluGlyGlyLysGlyPro 909
QY 699 TCAGGCTTCCAGCTGGCGCTTGTCCACAGGACA-----CCCCCA 661
Db 910 ArgGly-----GluThrGlyProAlaGlyArgProGlyGluValGlyProProGlyPro 928
QY 660 GGAATCTGGCAGTCTGCCACCATCGTCACAGGTCCTCCAGAACTCTCGCAGCTGCTCAACACAGG 601
Db 929 Gly-----ProAlaGlyGluLysGlySerProGlyAlaAasp----- 940
QY 600 GAAAGACAGGTTCCAGTCCATGCGGTGCTCCAGAACTCTCGCAGCTGCTCAACACAGG 541
Db 941 -----GlyProAlaGlyAlaProGlyThrProGlyProGluGlyLysAla 955
QY 540 AGCCGTAGCCAGCTTATCATTTATCAAAACCTCCGCGAGAATTTCTTGAAGAGGTGCCTC--- 484

```

```

:::
956 GlyGlnArgGlyValValGly-----LeuProGlyGlnArgGlyGluArgGlyPhePro 973
QY 483 -----GGTAGCTCATGG-----TCCGACAGAGCGGTGGAACTGATAAT 445
Db 974 GlyLeuProGlyProSerGlyGluProGlyLysGlnGlyProSerGlyAlaSerGlyGlu 993
QY 444 AAGACACACACAGATCCTTGGGGTTGCGAGCCATATAGATGACCTTGAGTCTCCATTGT 385
Db 994 ArgGlyProProGlyProMetGly-----ProProGlyLeu----- 1005
QY 384 GGAGGTCAGAGGGCAGAAACGGTAGGCGAGGTGGCTCTTATGAGCGGGGAGAGGTCA 325
Db 1006 AlaGlyProProGlyGluSerGlyArgGluGly----- 1016
QY 324 GTTCTTGTATGATGTCAGGCGCGGCTGTGGTACTCCAGGA-----CCGGGA 277
Db 1017 -----AlaProGly--AlaGluGlySerProGlyArgAaspGlySerProGly 1031
QY 276 GCTGCTCGATGTTTCATCAAGCCGATCTCATCGGGGTGAG----- 235
Db 1032 Ala-----LysGlyAaspArgGlyGluThrGlyProAlaGlyPro 1044
QY 234 -----CGCCTGCTCACCAGTAGACACACCTCTCTGCAGCAAGCTGGTCCGACTTGG 181
Db 1045 ProGlyAlaProGlyAlaProGlyAlaProGlyProValGlyProAlaGlyLysSer--- 1063
QY 180 GGTAGGTGACATCCACACGTCGCTGGCGCCGACCGGAGAGTTGGCGATCTCTCCATCT 121
Db 1064 -----GlyAaspArgGlyGluThrGlyProAlaGlyProAlaGlyProValGly 1079
QY 120 TCCGCGCGCAGAGGGCGGCGCAGCGCCATGCCAAGTACTGCTCTCGAACT 61
Db 1080 ProAlaGlyAlaArgGlyProAlaGlyProGlnGlyProArgGlyAaspLysGlyGluThr 1099
QY 61 ----- 61
Db 1100 GlyGluGlnGlyAaspArgGlyLysGlyLysGlyPheSerGlyLeuGlnGlyPro 1119
QY 60 -----CCCCGGGGTCTGGGGTCTCGG----- 37
Db 1120 ProGlyProGlySerProGlyGluGlnGlyProSerGlyAlaSerGlyProAlaGly 1139

RESULT 2
US-10-985-570-1
; Sequence 1, Application US/10985570
; Publication No. US20060100138A1
; GENERAL INFORMATION:
; APPLICANT: Olsen, David R.
; APPLICANT: Polarek, James W.
; APPLICANT: Yang, Chunlin
; TITLE OF INVENTION: IMPLANTABLE COLLAGEN COMPOSITIONS
; FILE REFERENCE: FP0305 US
; CURRENT APPLICATION NUMBER: US/10/985,570
; CURRENT FILING DATE: 2004-11-10
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 1466
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-985-570-1

Alignment Scores:
Pred. No.: 0,000101 Length: 1466
Score: 143.00 Matches: 76
Percent Similarity: 32.8% Conservative: 12
Best Local Similarity: 28.4% Mismatches: 106
Query Match: 3.2% Indels: 74

```

```
DB: 6 Gaps: 16
US-10-768-158-1 (1-2419) x US-10-985-570-1 (1-1466)
QY 765 TTCCCGCCGACCGGACGGGCTTCAGCTTGCAGCACTGCTCCACCAAGCT-----GGT 712
Db 172 TyrProGlyPro---AlaGlyPro-----ProGlyProProGlyProProGly 186
QY 711 GGCAGTGTCTCGCTCAGGCTTCCAGCTGGGCTTTGTACAGGACACCCCAAGG----- 658
Db 187 ThrSerGlyHisProGlySerProGlySerProGlyTyrGlnGlyProProGlyGluPro 206
QY 657 -----ATCTGGCAGCTGCTCCACCATCGTCACAGGTCCGATGCGATGCTTCATACT 604
Db 207 GlyGlnAlaGlyProSerGlyProGlyProGlyProGly----- 219
QY 603 TGAGAAAAGCACGTTGAGTCCATGCGGTCTCCAGAACTCTCTGACGGTCTCAAAACC 544
Db 220 -----AlaIleGlyProSerGlyProAlaGlyLysAspGly 231
QY 543 AGGAGCCGTAGCCAGCTTATCATTTATAAACCTCCGGCAGAAATTTTGAAGGTGCTC 484
Db 232 GluSerGlyArgPro-----GlyArgProGlyGluArgGlyLeu 244
QY 483 GGTAGCTCATGTCGCGCAGAGCGGTGGAACCTGATATAAGACACACCAAGATCCTTGG 424
Db 245 -----ProGlyProPro 248
QY 423 GGTTCGCGAG---CCATATAGATGACCT-----TGGAGTCTCCATTGTGGA--- 382
Db 249 GlyIleLysGlyProAlaGlyIleProGlyPheProGlyMetLysGlyHisArgGlyPhe 268
QY 381 ---GGTCAGAGGCGAGAAAGCGGTAGGCGAGGTGGCTCTTGATGAGGCGGGGAGAGGTCA 325
Db 269 AspGlyArgGlnGlyGluLysGlyGluThrGlyAlaProGlyLeuLysGlyGluAsnGly 288
QY 324 GTTCCT-----TGATGATGTCAGGCCCG----- 301
Db 289 LeuProGlyGluAsnGlyAlaProGlyProMetGlyProArgGlyAlaProGlyGluArg 308
QY 300 ---GCTGTGGTACTCCAGGACGGGAGCTGCTCGTGTGATGTTTCATCAAGCCGATCTCAT 244
Db 309 GlyArgProGlyLeuProGlyAlaAlaGlyAlaArgGlyAsnAspGlyAlaArgLysSer 328
QY 243 CGGGGTGAGCCCTCGCTCAGCAAGTAGACCACTCTCTGCAAGCTGCTGCGGACT 184
Db 329 AspGlyGln---ProGlyProProGly---ProProGlyThrAlaGlyPheProGlySer 346
QY 183 TGGGGTAGGTAGCATCCACACGTGCTGGGCGCCACCGGAAGTTGGCGATCTCTCCA 124
Db 347 ProGlyAlaLysGlyGluValGlyProAlaGlySerProGlySer---AsnGlyAlaPro 365
QY 123 TCTTCCCGCGCAGAAAGGCGGCGCCGACGCCATCGAAGTCTGCTCTCGA 64
Db 366 GlyGlnArgGlyGluProGlyProGlnGlyHisAlaGlyAlaGlnGlyProProGly--- 384
QY 63 ACTCCCGCGGGGTGCTGGGGGTCT 40
Db 385 ---ProProGlyIleAsnGlySer 391
```

RESULT 3

```
US-10-985-570-3
; Sequence 3, Application US/10985570
; Publication No. US20060100138A1
; GENERAL INFORMATION:
; APPLICANT: Olsen, David R.
; APPLICANT: Polarek, James W.
; APPLICANT: Yang, Chunlin
; TITLE OF INVENTION: IMPLANTABLE COLLAGEN COMPOSITIONS
; FILE REFERENCE: FP0305 US
; CURRENT APPLICATION NUMBER: US/10/985,570
; CURRENT FILING DATE: 2004-11-10
; NUMBER OF SEQ ID NOS: 3
```

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 3

; LENGTH: 1366

; TYPE: PR1

; ORGANISM: Homo sapiens

US-10-985-570-3

Alignment Scores:

Pred. No.:	0.000158	Length:	1366
Score:	140.50	Matches:	84
Percent Similarity:	28.5%	Conservative:	11
Best Local Similarity:	25.2%	Mismatches:	97
Query Match:	3.2%	Indels:	141
DB:	6	Gaps:	19

US-10-768-158-1 (1-2419) x US-10-985-570-3 (1-1366)

```
QY 840 CCATCTCTCTTTATACACCAAGTCAAACTTCTCATTCATGGAGACGGTGAAGATGTCCT 781
Db 204 ProGlyGluAsnGlyThrProGlyGlnThrGly----- 214
QY 780 TCCACAGCCCAACTCTTCCCGGCCACCG----- 751
Db 215 -----AlaArgGlyLeuProGlyGluArgGlyValGlyAlaProGlyProAlaGly 232
QY 750 -----GCAGGGCTTCAGCGTTGCAGC 730
Db 233 AlaArgGlySerAspGlySerValGlyProValGlyProAlaGlyProAsnGly---Ser 251
QY 729 ACTGCTCCACCA-----GCTGGT 712
Db 252 AlaGlyProProGlyPheProGlyAlaProGlyProLysGlyGluIleGlyAlaValGly 271
QY 711 GGCAGTCTCGTCAGGCTTCCAGCTGGGCTTGTACAGGACACCC---CCAGGAATC 655
Db 272 AsnAlaGlyProThrGly---ProAlaGlyProArgGlyGluValGlyLeuProGlyLeu 290
QY 654 ---TGGCCAGTGTCTCACCATGTCACAGGTCCCGATGCATGCTCTTCTACTTGAAG 598
Db 291 SerGlyProValGlyProProGlyAsnProGlyAlaAsnGly---LeuThrGlyAlaLys 309
QY 597 AAAGCAGCTTCGAGTCCA-----TGGCGTCTCCCAAGAACTCCT----- 559
Db 310 GlyAlaAlaGlyLeuProGlyValAlaGlyAlaProGlyLeuProGlyProArgGlyIle 329
QY 558 -----GCAGTGTCTCAACACGAGCGGTAGC 532
Db 330 ProGlyProProGlyAlaAlaGlyThrThrGlyAlaArgGlyLeuValGlyGluProGly 349
QY 531 CCAGCTTATCATTTCAAACTCCGCGCAGAAATTTCTTGAAGGTGCTCGGTAGTCATGG 472
Db 350 ProAlaGlySerLysGlyGluSerGlyAsn----- 359
QY 471 TCCGACAGAGCGGTGGAACCTGATAAAGACACACCAAGATCTTGGGTTGCGGCCA 412
Db 360 LysGlyLeuProGlySerAlaGlyProGlnGlyProProGlyProSerGlyGluGlu--- 378
QY 411 TATAGATGACCTTGGAGTCTCCATTGTGGAGGTCCAGAGGCGCAGAAAGCGGTAGGCGAGGT 352
Db 379 -----GlyLysArgGlyProAsnGlyGluAlaGly 388
QY 351 GGCTTTTGTAGCGCGGGGAGAGGTGATGTTCTTGTATGATGTCCAGGCCCGGCTGTGGGT 292
Db 389 -----SerAlaGly 391
QY 291 ACTCCAGGACCGGAGCTGCTCGTCGATGTTTCATCAAGCCGATCTCATCGGGGTACGGC 232
Db 392 ProProGlyProProGlyLeuArg----- 399
QY 231 CTGGGCTCACCA---AGTAGACCACTCTCTGACAGCAAGCTGTCGCGGACTTGGGGTAGG 175
Db 400 ---GlySerProGlySerArgGlyLeuProGlyAlaAsp-----GlyArg 413
```



```
Db 1975 GlyGluGlyArgLysGlyGlnGluArgSerArgThrHisLeuThrValProAlaGlySer 1994
Qy 498 -----CTTGAAGGTGCTCGTAGCTCATGTCGCGCAGAGCGGTGGAACTGATAAT 445
Db 1995 ThrGlnLeuProLeuCysProGlyLeuPheProCysGlyValAlaProGlyLeu---Cys 2013
Qy 444 AAGACACCA-----CCAGAT-----CCTTGGGGTTGCGAGCCATAT 409
Db 2014 LeuThrProGluGlnLeuCysAspGlyIleProAspCysProGlnGlyGluAsp----- 2031
Qy 408 AGATGACCTTGGAGTCTCATTTGTGGAGTCCAGAGGCGAGAAAGCGGT-----AGGSCA 355
Db 2032 -----GluLeuAspCysGlyGlyLeuProAlaLeuGlyGlyProAsnArgThr 2047
Qy 354 GGTGGCTTTGATGAGCGGGGAGAGGTGAGTCTCTTGATGATGTCCA----- 307
Db 2048 Gly-----LeuPro-----CysProGluTyrThrCys 2056
Qy 306 -----GGCCCGGCTGTG-----GGTACTCCAGGACCGGAGCT 274
Db 2057 ProAsnGlyThrCysIleGlyPheGlnLeuValCysAspGlyGlnProAspCysGlyArg 2076
Qy 273 GCTCGTCGATTTTCATCAAGCCGATCTCATCGGGTCAGCGCCCTGGCTCACCAAGTAGA 214
Db 2077 ProGlyGlnValGlyProSerProGluGlnGlyCysGlyAlaTrpGlyProTrpSer 2096
Qy 213 CCA---CCTCTCGCAGCAAGCTGTGTCGCGACTTGGGGTAGGTACCATCCACACAGT--- 160
Db 2097 ProTrpGlyProCysSerArgThrCysGlyProTrpGlyGlnGlyArgSerArgArgCys 2116
Qy 160 ----- 160
Db 2117 SerProLeuGlyLeuLeuValLeuGlnAsnCysProGlyProGluHisGlnSerGlnAla 2136
Qy 159 -----CGCTGGCGCGCACCGGGAGTGGCGA--- 133
Db 2137 CysPheThrAlaAlaCysProValAspGlyGluTrpSerThrTrpSerProTrpSerVal 2156
Qy 132 -----TCTCTCTCCATCTTCC 118
Db 2157 CysSerGluProCysArgGlyThrMetThrArgGlnArgGlnCysHisSerProGlnAsn 2176
Qy 117 CGCGCGAGAGGGCGGAGCC-----GCAGCCCATGGAACCTGGAAGTACT----- 73
Db 2177 GlyGlyArgThrCysAlaAlaLeuProGlyGlyLeuHisSerThrArgGlnThrLysPro 2196
Qy 72 -----TGC 70
Db 2197 CysProGlnAspGlyCysProAsnAlaThrCysSerGlyGluLeuMetPheGlnProCys 2216
Qy 69 TCTCGAACTCCC----- 58
Db 2217 AlaProCysProLeuThrCysAspAspIleSerGlyGlnValThrCysProProAspTrp 2236
Qy 57 -----CCGGGTGCTGG-----GGGTCTCG----- 37
Db 2237 ProCysGlySerProGlyCysTrpCysProGlyGlnValLeuGlySerGluGlyTrp 2256
Qy 36 -----CCTCGCTCTCCGCCATGC 19
Db 2257 CysValTrpProArgGlnCysProCys 2265
RESULT 6
US-10-985-570-3
; Sequence 3, Application US/10985570
; Publication No. US20060100138A1
; GENERAL INFORMATION:
; APPLICANT: Olsen, David R.
; APPLICANT: Polarek, James W.
; APPLICANT: Yang, Chunlin
; TITLE OF INVENTION: IMPLANTABLE COLLAGEN COMPOSITIONS
; FILE REFERENCE: FP0305 US
; CURRENT APPLICATION NUMBER: US/10/985.570
```

```
; CURRENT FILING DATE: 2004-11-10
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 1366
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-985-570-3
Alignment Scores:
Pred. No.: 0.0174 Length: 1366
Score: 116.00 Matches: 80
Percent Similarity: 31.2% Conservative: 13
Best Local Similarity: 26.8% Mismatches: 78
Query Match: 2.6% Indels: 127
DB: 6 Gaps: 20
US-10-768-158-1 (1-2419) x US-10-985-570-3 (1-1366)
Qy 43 CCCCCAGCACCCGGGGAGTTTCGAGAGCAAGT-----ACTTCGAGTTCCATGGCGTGC 96
Db 293 ProValGlyProProGlyAsnProGlyAlaAsnGlyLeuThrGlyAlaLysGlyAlaAla 312
Qy 97 GGCTGCCGCCCTTCTCGCGGGGAGATGGAGGAGATGCCCACTTCCCGTGGCGGCCA 156
Db 313 GlyLeu---ProGlyValAlaGly-----AlaPro 321
Qy 157 GCGACGTGTGATCGTCCACCTACCCCAAGTCCGGCACCCAGCTTCTGTCAGGA----- 208
Db 322 GlyLeuProGlyProArgGlyIleProGlyPro---ProGly-AlaAlaGlyThrThrGl 340
Qy 209 ----GGTGGTCTACTTGGTGAGCCAGGCGCTGACCCCGATGAGATCGGTTGATGAACA 264
Db 340 yAlaArgGlyLeuValGlyGluProGly-----ProAlaGlySerLysGlyGluSerGl 358
Qy 265 TCGACGAGCAGCTCCCGTCTCTGGAGTACCCACAGCGGGCTGACATCATCAAGGAC 324
Db 358 yAsnLysGlyGluProGlySerAlaGlyProGlnGlyProProGlyProSerGlyGluGl 378
Qy 325 TGACCTCTCCCGCTCATCAAGAG-----CCACCTGCCCTACCGCTTTC 369
Db 378 uGlyLysArgGlyProAsnGlyGluAlaGlySerAlaGlyProProGly----- 394
Qy 370 TGCCCTCTGACCTCCACAATGGAGACTCCAAGGTCTATATATGGCTCGCAACCCCAAGG 429
Db 395 -----ProGlyLeuArg-----Gl 400
Qy 430 ATCTGGTGGTGTCTTATTATTCAGTTCCACCGCTCTCTGCGGACCATGAGTACCGAGCA 489
Db 400 ySerProGly----- 403
Qy 490 CTTTTCAGAAATTCGCGGAGGTTTATGAATGATAAGCTGGGCTACGGCTCCTGTTTG 549
Db 404 ----SerArgGlyLeuProGlyAlaAsp----- 411
Qy 550 AGCAGCTGCAGGAGTTCTGGGAGCACCGCATGAGTCTGAACGTCCTTTTCTCAAGTATG 609
Db 412 -GlyArgAlaGlyValMetGlyProProGlySer----- 422
Qy 610 AAGCATGCTATCGGGACCT-----GGTGACGATGGT----- 640
Db 423 -ArgGlyAlaSerGlyProAlaGlyValArgGlyProAsnGlyAspAlaGlyArgProGl 442
Qy 641 -----GGAGCAGCTCGCCAGATT---CCTGG 663
Db 442 yGluProGlyLeuMetGlyProArgGlyLeuProGlySerProGlyAsnIleGlyProAl 462
Qy 664 GGGTGTCTCTGTGACAAGCCCGCAGCTGGAAGCCCT-----GACGAGCAGCTGCCACGAGC 717
Db 462 aGlyLys-----GluGlyProValGlyLeuProGlyIleAspGlyArgProGlyProIl 480
Qy 718 TGGT-----GGACCAGTGTGTGCA 735
```



```
Db      480 eGlyProValGlyAlaArgGlyGluProGlyAsnIleGlyPheProGlyProLys----- 498
Qy      736 ACGCTGAGGCGCTCCCGTGGCGGGAAGTTGGGCTGTGAAGACAT 787
Db      499 -----GlyProThrGlyAspProGlyLysAsn---GlyAspLysGlyHis 512

RESULT 7
US-10-196-749-525
; Sequence 525, Application US/10196749
; Publication No. US20060094864A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C340
; CURRENT APPLICATION NUMBER: US/10/196,749
; CURRENT FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 525
; LENGTH: 4440
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-196-749-525

Alignment Scores:
Pred. No.: 0.0334 Length: 4440
Score: 115.00 Matches: 86
Percent Similarity: 37.3% Conservative: 50
Best Local Similarity: 23.6% Mismatches: 181
Query Match: 2.6% Indels: 48
DB: 6 Gaps: 10

US-10-768-158-1 (1-21419) x US-10-196-749-525 (1-4440)
Qy      58 GGGAGTTCGAGAGCAAGTACTTCGAGTTCATGGCGTGGCGCTTCTGCGCGC 117
Db      3325 GlyThrCysAlaAlaThrGlyThrCysCysThrCysThrCysAlaThrAlaThr 3344
Qy      118 GGAAGATCGAGAGATCGCCAACTTCCCGTGGCGGCCAGGACGTGTGGATCGTCACT 177
Db      3345 GlyThrThrAlaAlaCysThrThrThrCysThrThrAlaThrThrGlyAlaThrThr 3364
```

```
Qy      178 ACCCCAGTCCGGCACCAGCTTGTCTGAGGAGGTGTTCTACTTGGTGAGCCAGGCGCTG 237
Db      3365 ThrGlyThrThrAlaGlyAlaGlyCysThrThrThrGlyThrAlaThrAlaThr 3384
Qy      238 ACCCCGATGAGATCGGCTTGATGAACATCG-----ACGAGCAGCTCCCGCTCCTGG 288
Db      3385 ThrGlyThrGlyAlaAlaThrAlaThrAlaGlyCysThrThrThrGlyThrGly 3404
Qy      289 AGTACCCACAGCGCGGCTGGACATCATCAAGNACTGACCTCTCCCGGCTCATCAAGA 348
Db      3405 AlaThrCysAlaGlyAlaAlaThrThrAlaThrGly-----GlyAlaAlaThrGly 3422
Qy      349 GCCACCTGCCCTACCGCTTCTGCTCCTCTGACCTCCACATGGAGACTCCAAAGGTCTCT 408
Db      3423 ThrThrGlyThrThrGlyThrCys-----ThrThrThrAlaGlyThrThrThrGly 3440
Qy      409 ATATGGCTCGCAACCCCAAGGATCTGGTGGTGTCTTATTATCATGTTCCACCGCTCTCTGC 468
Db      3441 CysThrThrAlaThr-----CysAlaAlaGlyThrThrThrAlaAlaAla 3455
Qy      469 GGACCATGAGTACCGAG-----GCACCTTCAAGATTTCTGCGGAGGTTTATGA 519
Db      3456 GlyCysCysAlaThrThrCysAlaGlyAlaGlyAlaThrGlyThrGlyThrAlaAla 3475
Qy      520 ATGATAAGCTGGGCTACGGCTCCTGGTTGAGCAGCTGCAGGAGTCTCGGAGCACCGCA 579
Db      3476 AlaThrGlyThrGlyThrAlaThrGlyThrGlyThrThrAlaAlaAlaThrThrThr 3495
Qy      580 TGGACTCGAACAGTCTTTTCTCAAGTATGAAGACATGCATCGGACCTCGTGACGATGG 639
Db      3496 AlaThrGlyThrCysThrThrAlaThrThrThrGlyThrGlyThr----- 3511
Qy      640 TGGAGCAGCTGGCGAGATTCTGGGGGTGCTCTGTGACAAGGCCACGCTGGAAGCCCTGA 699
Db      3512 ThrThrCysThrGlyAlaThrGlyCysThrCysAlaThrAlaThrGlyThrThrThr 3531
Qy      700 CGGAGCACTGCCACCAGCTGGTGGACCAAGTGTGCAACGCTGAGGCCCTGCCGCGGCC 759
Db      3532 AlaGlyAlaAlaThrGlyThrCysAlaAlaGlyCysAlaAlaThrCysCysAla 3551
Qy      760 GGGGAAGAGTTGGGCTGTGGAAGGACATCTTCAACCGTCTCCATGAATGAGAAGTTGACT 819
Db      3552 AlaCysThrThrAlaThrGlyAlaThrThrAlaCysAlaThrAlaAlaAlaThrAlaThr 3571
Qy      820 TGGTGTATAAACAGAACATGGGAAGTGTGACCTCAGCTCAGTTTGACTTTA----- 868
Db      3572 ThrCysThrCysCysCysAlaThrAlaThrThrGlyCysThr-CysThrThrAlaGly 3591
Qy      869 -----TTTATAATAACAGAACATGGGAAGTGTGACCTCAGCTCAGTTTGACTTTA----- 868
Db      3591 yCysAlaThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 3611
Qy      922 TAGCCAAAAGTCTGTATGCATTTTATTTCTTGTGTCGACAAACT----- 968
Db      3611 rAlaAlaCysCysCysThrThrThrGlyThrThrCysThrThr-ThrCysAlaGlyGlyA 3631
Qy      969 -----CTGGAGCAGCGTGTGAACACAGCGGGGGAAGGGAAGAGC----- 1007
Db      3631 laAlaThrThrAlaCysThrThrThrGlyCysGlyThrAlaAlaGlyAlaAlaA 3651
Qy      1008 -----GGCGTGAGCGGAGGAGTGATGATGATTTCCCAACCGAAGCAGCTGTCTCG 1056
Db      3651 laThrGlyGlyCysThrGlyCys-----ThrThrCysCysA 3665

RESULT 8
US-10-985-570-1
; Sequence 1, Application US/10985570
; Publication No. US20060100138A1
; GENERAL INFORMATION:
```



```
QY 339 GCGCGGAGAGGTGAGTTCCTTGATGATG-----CGAGGCCGGCTGTG 295
D 609 ArgGlyProLeuGlyLeuProGlyAlaSerGlyLeuAspGlyArgProGlyProPro--- 627
QY 294 GGTACTCAGGACCGGAGCTGCTCGTGAATGTTTCATCAAGCCGATCTCATCGGGGTGAG 235
D 628 GlyThrProGlyProLeuGlyValProGlyProAlaGlyProLysGlyGluArgGlySer 647
QY 234 CGCCTGGCTCACAAGTAGACCACTCTCTCCAGCAAGCTGGTCCGAGCTTGGGGTAGG 175
D 648 Lys---GlyAspProGlyMetThrGlyPro----- 656
QY 174 TGACGATCCACACGTGCTGGCGCGCACCGGAGTGGCGATCTCTCCATCTTCCCGC 115
D 657 -----ThrGlyAlaAlaGlyLeuProGly---LeuHisGlyProProGlyAspLys 672
QY 114 GCGAAGGGCGGCGAGCGCGACGCGCATCGAAGTCTGAACTCGAAGTACTTGTCTCGAATCCCGC 55
D 673 GlyAsnArgGlyGluArgGlyValGlyGlySerArgGly-----ProLysGlyAspLys 690
QY 54 GGGTGTGGGGTCTCGGCGCTCGCTCTCCGCCATGC 19
D 691 GlyAspGlnGlyAlaProGlyLeuAspAlaProCys 702

RESULT 10
US-10-511-937-2534
; Sequence 2534, Application US/10511937
; Publication No. US2006008836A1
; GENERAL INFORMATION:
; APPLICANT: EXPRESSION DIAGNOSTICS, INC.
; APPLICANT: Wohlgenuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; APPLICANT: Prentice, James
; APPLICANT: Morris, MacDonald
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND MONITORING TRANSPLANT REJECTION
; FILE REFERENCE: 50661200104
; CURRENT APPLICATION NUMBER: US/10/511,937
; CURRENT FILING DATE: 2004-10-19
; PRIOR APPLICATION NUMBER: PCT/US2003/012946
; PRIOR FILING DATE: 2003-04-24
; PRIOR FILING DATE: 2002-04-24
; PRIOR FILING DATE: 2002-04-24
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 3117
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2534
; LENGTH: 529
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-511-937-2534

Alignment Scores:
Pred. No.: 0.0901 Length: 529
Score: 105.50 Matches: 47
Percent Similarity: 37.4% Conservative: 18
Best Local Similarity: 27.0% Mismatches: 56
Query Match: 2.4% Indels: 43
DB: 6 Gaps: 8

US-10-768-158-1 (1-2419) x US-10-511-937-2534 (1-529)
QY 852 GGTACACTTCCCATCTTCTGTTTATACACCAAGTCAAACTTCTCATTCATGGAGCGG 793
D 348 GlyLysTyrPheSerSerAsnProIlePheProSerGlnSerAlaAlaSer----- 364
QY 792 TGAAGATGCTCTTCACAGCCCACTCTTCCCGGCCACCGGCGGCGCTCAGGCTTGC 733
D 733
```

```
D 365 -----SerAlaIleAlaPheGlnProValGlyThrGlyGlyValGlnLeuCys 380
QY 732 AGCACTGCTCCACAGAGTGGCAGTGTCCGTCAGGCTTCCAGCTGGGCCCTGTGTAC 673
D 381 -----GlyGlyGlySerThrGlySer-----LysGlyProCysSer 392
QY 672 AGGACACCCCGAGAAATCTGGCAGCTGCTCCACCATCGTCACCAAGTCCCGATGCAATG 613
D 393 ProSerSerSerArgVal---ProSerSerSerSerSerSerSerSerSerSerSerPro 411
QY 612 CTTCATACTTACAGAAAAGACACGTTCCAGTCCATGCGGTGCTCCCAAGACTCTCGACGT 553
D 412 TyrHisProCysGlySerAlaSerGlnSerProCysSerProGlyThrGlySerPhe 431
QY 552 GCTCAACACAGGAGCGGTAGCCAGCTTATCATTAACCTCCGCGCAGAAATCTTGAA 493
D 432 SerSerSerSerSer-----SerGlnSerSerGlyLysIleLeu 445
QY 492 AGGTGCTCGGTAGCTCATGTCGCGCAGAG----- 463
D 446 GlnProCysGlySerLysSerSerSerSerGlyHisProCysMetSerValSerSerLeu 465
QY 462 -----AGCGGTGGAAGTATAATAAGACACCAAGTCTCT---TGGGGTTGGAGCCA 412
D 466 ThrLeuThrGlyGlyProAspGlySerProHisProAspProSerAlaGlyAlaLysPro 485
QY 411 TATAGATGACCTTGGAGTCTCCATTGTGGAGGTGAGAGGCA 370
D 486 -----CysGlySerSerSerAla 491

RESULT 11
US-10-505-928-493
; Sequence 493, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 493
; LENGTH: 744
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-493

Alignment Scores:
Pred. No.: 0.103 Length: 744
Score: 105.50 Matches: 77
Percent Similarity: 32.4% Conservative: 16
Best Local Similarity: 26.8% Mismatches: 73
Query Match: 2.4% Indels: 121
DB: 6 Gaps: 20

US-10-768-158-1 (1-2419) x US-10-505-928-493 (1-744)
QY 26 GGAGAGCGAGCGCCGAGAC-----CCCCAGCACCCCGGGGAGTTGAGAG 70
D 316 GlyLysProGlyGlnAspGlyIleProGlyGlnProGlyPheProGlyGlyLysGlyGlu 335
QY 71 CAAGTACTTCGAGTTCCATGCGGTGCGGCTGCCGCCCTTCTGCGCGGGAAGATGGAGGA 130
D 336 GlnGlyLeu-----ProGlyLeuProGlyProGly--- 346
QY 131 GATCGCCAACTTCCGGTGGCGCCAGCGAGTGTGGATCGTCACTACCCCAAGTCCGG 190
D 347 -----LeuProGly----- 349
QY 191 CACCAGCTTGTGCGAGGAGTGTCTACTTGTGTAGCCAGGCGGTGACCCCGATGAGAT 250
```

```

Db      350  -----:||||:|||||
Qy      251  CGGCTTGATGAACATCGACGAGCTCCCGTCTCT-----GGA 289
Db      355  -----PheProGlyProLysGlyAspArgGlyMetGlyGly 366
Qy      290  GTACCCACAGCC---GGGCTGGACATCATCAAGAACTGACCTCTCCCG-----337
Db      367  ValProGlyAlaLeuGlyProArgGlyGluLysGlyProIleGlyAlaProGlyIleGly 386
Qy      338  ---CCTCATCAAGAGCCA---CCTGCCCTPACCGCTTCTGCCCTCTGACCTCCACAA 388
Db      387  GlyProProGlyGluProGlyLeuProGlyIlePro---GlyProMetGlyProPro---404
Qy      389  TGGAGACTCAAGTCACTATATGGCTCGCAACCCCAAGGATCTGGTGGTCTCTTA---445
Db      405  -----GlyAlaIleGlyPheProGlyProLysGlyGluGlyGlyIleValGly 420
Qy      446  -----TTATCAGTTCCACCGCTCTCTGCCGAGCATGAGCTACCG 484
Db      421  ProGlnGlyProProGlyProLysGlyGluProGlyLeuGlnGlyPheProGlyLysPro 440
Qy      485  AGGCACCTTTCAAGAACTCTCCGAGGTTTATGAATGATGAAGCTGGCTACGGCTCCTG 544
Db      441  GlyPheLeuGlyGluValGlyPro---:|||||ProGlyMetArgGlyLeu 454
Qy      545  -----GTTTGAGCAGCTGCAGGAGTTCTGGGAGCACCCGCACTGGACTCGAA 589
Db      455  ProGlyProIleGlyProLysGlyGluAlaGlyGlnLysGlyValPro---GlyLeu---472
Qy      590  CGTGCTTTTCTCAAGTATGAAGACATGCATCGGAGCTCGGACCTGGTGAGCATGGT-----GGA 643
Db      473  -----ProGlyValProGlyLeuLeuGly 480
Qy      644  CGAGCTGCCAG-----ATTCTGGGGTGTCTGTGACAAAGGCCA-----685
Db      481  ProLysGlyGluProGlyIleProGlyAspGlnGlyLeuGlnGlyProProGlyIlePro 500
Qy      686  ---GCTGGAAGCCCTGACGCGAGCACTGCCAGCTGGTGGACCAAGTCTGCAACGCTGA 742
Db      501  GlyIleGlyGlyProSerGlyProIleGlyProProGlyIlePro-----515
Qy      743  GGCCCTGCCCGTGGCGGGG 763
Db      516  GlyProLysGlyGluProGly 522

```

RESULT 12

```

US-10-505-928-799
; Sequence 799, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; PRIOR FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 799
; LENGTH: 1384
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-799

```

```

Alignment Scores:
Pred. No.:      0.131      Length:      1384
Score:          105.50     Matches:      109
Percent Similarity: 32.4%   Conservative: 45
Best Local Similarity: 22.9% Mismatches: 174
Query Match:      2.4%     Indels:      148

```

```

DB:      6      Gaps:      27
US-10-768-158-1 (1-2419) x US-10-505-928-799 (1-1384)
Qy      1140  CTTTATCTTACGGTCCAGGCCATTGGAACTGCAATGTGGAGACTGTT-----1093
Db      103  LeuTyProCysGlyAsnAla-----CysThrIleAspGlyLeuProValArgGln 119
Qy      1092  TGTATATCAGACATGGAGAGGCTGCACAGTCTTAAAGCGCAGACAGCTGTCGGT-----1039
Db      120  ProThrArgLeuThrGlnGlyCysMetLeu-----CysLeuGlyGlnSer 134
Qy      1038  -----TGGGAATCATCATCATCTCCCTCCGCG 1015
Db      135  ThrPheLeuArgPheAsnHisProAlaGluAlaLysTrpMetLysSerMetIleProAla 154
Qy      1014  TCAGCGCGCTCTCCCTTCCCGCGCTGTT-----985
Db      155  GlyGlyArgAlaProGlyProTySerProValProAlaGluSerGluSerLeuVal 174
Qy      984  -----CACAGCTGCTTCCAGAGTTTGTCCAGCAAGAAATAAATGAATGCATACAG 934
Db      175  AsnGlyAsnHisThr-----ProGlnThrAlaThrArgGlyProSerAlaCysAlaSer 192
Qy      933  GACTTTTGGCTAGTAGACTGTCTGGGTATTGTGACATGCAGGTTGTGTTTCTGTTATT 874
Db      193  HisSerSerLeuValSerSerIleGluLysAspLeuGlnGluIleMetAspSerLeuVal 212
Qy      873  ATAAATAAAGTCAAAAGTGGAGTGCACACTTTCCTCTCTGTTTATACACCAAGTCAA 814
Db      213  LeuGluGluProGlyAlaAlaGlyLys-----LysProAlaAla 225
Qy      813  ACTTCTCATTCATCGAGACGGTGAAGATGTCTTCCACAGCCCAACTCTTCCCC---GGC 757
Db      226  ThrSer-----ProLeuSer 230
Qy      756  CCAGCGGACGGGCTCAGCGTTGCAGCACTGGTCCACCAGCTGGTGGCAGTCTCCGTCA 697
Db      231  ProMetAlaAsnGlyArgTyLeuSerProPro-----ThrSerProGly 247
Qy      696  GGGCTTCAGCTGGGCTTGTGCAC---AGGACACCCCGAGGAATCTGCCAGCTGCTCCA 640
Db      248  AlaMetSerValGlySerSerTyGluAsnThrSerProAlaPheSerProLeuSerSer 267
Qy      639  CCATCGTCACAGGTCGCCGATGTCATGTTTCATATCTTGAGAAAAAGCAGCTTCGAGTCCA 580
Db      268  ProAlaSerSerGlySerCysAlaSer---HisSerProSerGlyGlnGluProGlyPro 286
Qy      579  TGGCGTCTCCAGAACTCTCGCAGCTGCTCAAAACAGGAGCGGTAGCCAGCTTATCAT 520
Db      287  SerValProProLeuValProAla-----ArgSerSerSerTyHis 300
Qy      519  -----TCATAAACCTCCGCGCAGAAATCTTGAAGGTGCTCGGT 481
Db      301  LeuAlaLeuGlnProProGlnSerArgProSerGlyAlaArgSerGlu-----316
Qy      480  AGCTCATGTCGCCGACAGAGCGGTGGAACCTGATTAATAAGCACACCAAGCTTGGGGT 421
Db      317  SerProArgLeuSerArgLysGlyGly-----HisGluArgProProSerPro-GlyLe 334
Qy      420  TGGCAGGCATATAGATGACCTTGGAGTCTCATTTGTG-----GAGGTCAAG-----375
Db      334  uArgGlyLeuLeuThrAspSerProAlaAlaThrValLeuAlaGluAlaArgAlaTh 354
Qy      374  -----GGGCAGAAAGCGGTAGGCGAGGTGCTCTTGTATGAGGC---337
Db      354  rGluSerProArgLeuGlyGlyGlnLeuProValValAlaIleSerLeuSerGluTyPr 374
Qy      336  -----GGGAGAGGTGTCAGTCTTGTGATGTCAGGC---CCGGCTGGGTACTCC 287
Db      374  oAlaSerGlyAlaLeuSerGlnProThrSerIleProGlySerProLysPheGlnProPr 394
Qy      286  AGGACCGGAGCTGCTCGTCGATGTTTCATCAAGCCGATCTTCATCGGGGTGAGCCCTGG 227

```


Db 1042 Ser-----ProSerThrSerProProSerAlaLeu--- 1051
Qy 2009 GGCTACTGCCCCCG-----ACTCTCCATCGCTTCAGCGAATGTGGG 1968
Db 1052 -----ProProAspValAspProGlyGluProMetHisAlaSerProPro----- 1066
Qy 1967 TGAGAAATAGGAGTGAAGCAATAAGACCTCACCTTTACTCCCTCAGATACACGAG 1908
Db 1067 -----ArgLysAlaArgValArgLysAlaSerSerTyrTyrProGlu 1080
Qy 1907 GGATGAGTGACAGACACG-----AGGGCGCGCTGCCAACCTGATCCCGAGTAA 1857
Db 1081 GlyAspGlyAspSerSerHisLeuProAlaLysLysSerCysAspGluAsp----- 1097
Qy 1856 ACAGGGCTACACTCGCAAAATGTCTCCCGCTGAGCGCTCAGCGCTCAGGTGCACAGAACAA 1797
Db 1098 -----LeuThrSerGlnLysValAspGluLeuGlyGlnLysArgArg----- 1111
Qy 1796 ACTGAATGATAAGACCC-----GGGATTTTAGAAATTTACTATTAAAGCATCTGC 1746
Db 1112 -----ArgProLysSerAlaProSerPheArgProLysLeuAlaProValVal 1127
Qy 1745 TTGCCAGCATTAGTGGCTATCACTTCAGGAAACAAAGTAAGGCTGTGGAATTCAGCTCA 1686
Db 1128 IleProAla-----GlnPheLeuGluGlnLysCysValProAlaSerGly 1143
Qy 1685 AATTCGAGCGCAACTCGCTCTTTGGTGGCG-----ACTCTCCAC 1644
Db 1144 GluLeuSerProGluLeuGlnLysAlaProTyrSerProGlyHisSerSerArgHis 1163
Qy 1643 -----CACTGCTGCAATGCTCTGACTGCTATGACTTACTCCACCTGCCAGCTC 1593
Db 1164 SerAsnProProLeuTyrProSerArgProSer---ValGlyThrValProArgSerLeu 1182
Qy 1592 ACCAGGCAGACTTCCTCCAGCAC-ATGCTCCGCGGGCTCCTCTGC----- 1549
Db 1183 ThrProSerThrThrValSerSerIleLeuArgAsnProIleTyrThrValArgSerHis 1202
Qy 1548 ---ATTGGCCTTTG---GAGCGCGCCCTCGA----- 1522
Db 1203 ArgValGlyProCysSerSerProProAlaAlaArgAspAlaGlyProGlnGlyLeuHis 1222
Qy 1521 ---AGTCTCTGCAGAGGAGCTCCATCTCTGAAGACCTAGTGAACAATTTCTGTCTCA--- 1468
Db 1223 ProSerValGlnHisGlnGlyArgLeuSerLeuAspLeuSerHisArgThrCysSerAsp 1242
Qy 1467 TTTCAGGAA----- 1459
Db 1243 TyrSerGluMetArgAlaThrHisGlySerAsnSerLeuProSerSerAlaArgLeuGly 1262
Qy 1459 ----- 1459
Db 1263 SerSerSerAsnLeuGlnPheLysAlaGluArgIleLysIleProSerThrProArgTyr 1282
Qy 1458 -----GGTCTGGAAGTGAAGTGAAGTCCGTCAGATATAACAAAGTCCAGAT 1414
Db 1283 ProArgSerValValGlySerGluArg---GlySerValSerHisSerGluCys-SerTh 1301
Qy 1413 TCCCCCAACAGACCACTATCTCTGCTCTCCACCACTATCTCAGAACCAAGTGAGAA 1354
Db 1301 rProProGlnSerProLeuAsnIleAspThrLeuSerSerCysSerGlnSerGlnThrSe 1321
Qy 1353 CAGAAATGCTGTGCTTATAGTA----- 1332
Db 1321 rAlaSerThrLeuProArgIleAlaValAsnProAlaSerLeuGlyGluArgArgLysAs 1341
Qy 1331 -----ATAAACTACAAAATAAAGATTAAACCTTGA 1300
Db 1341 pArgProTyrValGluGluProArgHisValLysValGlnLys----- 1355
Qy 1299 GATCTTCTACCTTCGGGTGTGGCTCGCTCTGATTCCTCCCTTGGAATGAATTTTATT 1240
Db 1299 ----- 1240

Db 1356 -----GlySer-----GluProLeuGlyIleSerIleValSe 1366
Qy 1239 TGGTTTACTGACATTTATGTAGATTTCAGTGAAGAGCTCTATATAATACAAATACGG 1180
Db 1366 rGly----- 1367
Qy 1179 GGTGAAAAGCAGACACATTTCTAGTTGCATATATTACAGGGCTTTATCTTACGGTCCAGGC 1120
Db 1368 -----GluLysGlyGlyIleTyrValSerLysValThrValGlySerIleAlaHisGlnAl 1386
Qy 1119 CATTGGAACCTGCAATGTGGAGACTGTTTGTAAATCAGACATGGAGAGGCTGCACGTTCTAA 1060
Db 1386 a---Gly-LeuGluTyrGlyAsp----- 1392
Qy 1059 AGGCGACAGACTGCTCTCGGTGGGAATCATCACACTCCCTCCGCTCAGCGCTCTTCC 1000
Db 1393 -----GlnLeuLeuGluPheAsnGlyIleAsnLeuArgSerAlaThrGluGlnGlnAl 1410
Qy 999 CTTCCCGCGCTGTTTTCACAGCTCTTCCAGAGTTTCTCCAGCAAGGAATAAATGAATGC 940
Db 1410 laArgLeuIleIleGlyGlnGlnCysAspThrIleThrIleLeuAlaGlnTyrAsnProH 1430
Qy 939 ATACAGGACTTTTGGCT-----AGTAGACTGTCTGGGTATTGTGACATGCAGGTGTT 886
Db 1430 isValHisGlnLeuSerSerHisSerArgSerSerHisLeuAspProAlaGly-Thr 1449
Qy 885 GTTCTCTGTTATTATAATAAAGTCAAACTGAGGTGCACACTTCCCATCT-----TCT 832
Db 1450 HisSerThrLeu-----GlnGlySerGlyThrThrThrProGluHisProSer 1465
Qy 831 GTTTTATACCAAGTCAAACTTCTCATTCATGGAGAGCGTGAAGATGCTCTCCACAGCC 772
Db 1466 ValIleAspProLeuMetGluGlnAspGlu----- 1475
Qy 771 CAACTCTTCCCGCGCCACGCGGCTCAGCGTTCGACACTGTGTCCACAGCTGT 712
Db 1476 -----GlyProSerThrProProAlaLysGlnSerSerArgIleAlaGly 1491
Qy 711 GGCAGT-----GCTCGCTCAGGCTTCCA 688
Db 1492 AspAlaAsnLysLysThrLeuGluProArgValValPheIleLysLysSerGlnLeuGlu 1511
Qy 687 GCTGGGCT-----TGTCACAGGACACCCCGCAAGATCTGCCAGCT--GCTCCACCA 637
Db 1512 LeuGlyValHisLeuCysGlyGlyAsnLeuHisGlyValPheValAlaGluValGluAsp 1531
Qy 636 TCCTCACCA-----GCTCCGATGCTGT-----CTTCACTATTGAG----- 600
Db 1532 AspSerProAlaLysGlyProAspGlyLeuValProGlyAspLeu-IleLeuGluTyrGl 1551
Qy 599 -AAAAAGCAGCTTCGAGTCCATGCGTCTCCAGAACTCTCGACGTCGTCACCAACGAGG 541
Db 1551 ySerLeuAspValArgAsnLysThrValGluGluValTyrValGluMetLeuLysProAr 1571
Qy 540 AGCGGTACCGCAGCTTATCATTAACCTCCGCGAGAAATCTTGAAGGTGCTCGCT 481
Db 1571 gAspGlyValArgLeuLysValGlnTyrArgProGluGluPheThrLysAlaLysGlyLe 1591
Qy 480 AGCTCATGTCTCCGAGAGAGCGGTGGAACCTGATATAAGACACCAACGAGTCTTGGGGT 421
Db 1591 u-----ProGlyAspSerPheTyrIleArgAlaLeuTyrAspArgLeuAlaAspVa 1608
Qy 420 TGCAGGCCATATA-----GATGACTTGGAGTCTCCATTGTGGAGGTGCAGAGG 373
Db 1608 lGluGlnGluLeuSerPheLysLysAspAspIle---LeuTyrValAspAspThrLeuPr 1627
Qy 372 GCAGAAAGCGGTAGCGAGGTGGCTCTTG-----ATGA 340
Db 1627 oGlnGlyThrPheGlySer-TripMetAlaTrpGlnLeuAspGluAsnAlaGlnLysIleG 1647
Qy 339 GCGGGGAGAGGTGAGTTCC-----TTGATGA 313
Db 1647 lnArgGlyGlnIleProSerLysTyrValMetAspGlnGluPheSerArgArgLeuSerM 1667


```
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-196-749-310

Alignment Scores:
  Pred. No.: 0.203      Length: 596
  Score: 101.50      Matches: 110
  Percent Similarity: 36.7%      Conservative: 49
  Best Local Similarity: 25.4%      Mismatches: 224
  Query Match: 2.3%      Indels: 50
  DB: 6      Gaps: 15

US-10-768-158-1 (1-2419) x US-10-196-749-310 (1-596)

Qy 61 AGTTCGAGAGCAAGTACT-----TCGAGTTCCATGGCGTGGCGTGGCGC 105
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 159 SerSerArgAlaSerThrAlaThrAsnSerGluSerSerThrLeuSerSerGlyAla--- 177
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 106 CTTCTCGCGGGAAGATGAGGAGATCGCAACTTCCCGTGGCGGCCCGACGCTGT 165
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 178 SerThrAlaThrAsnSerAspSerSerThrThrSerSerGlyAlaSerThrAlaThrAsn 197
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 166 GGATCGTCACCTACCCCAAGT-----CCGGCACCAGCTTCTGTCAGGAGGTGTCTACT 219
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 198 SerGluSerSerThrThrSerSerGlyAlaSerThrAlaThrAsnSerGluSerSerThr 217
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 220 TGGTGAGCCAGGCGCTGACCCGATGATGATGCGCTTGATGAACATGACGAGCGTCC 279
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 218 ValSerSerArgAlaSerThrAlaThrAsnSerGlu---SerSerThrThrSerSerGly 236
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 280 CGGTCTCGAGTACCCAGCGCGGCTGGACATCATCAAGAACTGACCTCTCCCGCC 339
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 237 AlaSerThrAlaThrAsnSerGluSerArgThrThrSerAsnGlyAlaGlyThrAlaThr 256
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 340 TCATCAAGAGCC-----ACCTGCCCTACCGCTTCTGCTCCCTCCTCACCAATGGAG 393
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 257 AnSerGluSerSerThrThrSerSerGlyAlaSerThrAlaThrAsnSerAspSerSer 276
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 394 ACTCCAAGTCTATATGCTGCTGCAACCCCAAGGATCTGGTGGTGTCTTATATCATGT 453
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 277 ThrValSerSerGlyAlaSerThrAlaThrAsnSer-----GluSer 290
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 454 TCCACCGCTCTCTGGGA---CCATGAGCTACCGAGCGACCTTTCAGAAATCTGCCGGA 510
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 291 SerThrThrSerSerGlyAlaSerThrAlaThrAsnSerGluSerSerThrThrSerSer 310
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 511 GGTATTATGAATGATAAGCTGGCTACGGCTCTCTGTTTGAGCACGTGCAGGAGTCTGGG 570
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 311 Gly-----AlaSerThrAlaThrAsnSerAspSerSerThrThrSerSerGlyAla 327
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 571 AGCACCGCATGGACTCGAACGTGCTTTTCTCAAGATGATGAACATGATCGGACCTGG 630
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 328 GlyThrAlaThrAsnSerGluSer-----SerThrValSerSerGlyLeuSerThrVal 345
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 631 TGACGATGGTGAGCAGCTGGCCAGATTCCTGGGGGTCTCTGTGACAGGCCCGACGTGG 690
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 346 ThrAsnSerGluSerSerThrProSerSerGlyAlaAlaThrAlaThrAsnSerGluSer 365
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 691 AAGCCCTCAGCGAGC-----ACTGCCACC-----AGCTGG 720
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 366 SerThrThrSerSerGlyAlaAsnThrAlaThrAsnSerGluSerSerThrValSerSer 385
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 721 TGGACCATGCTGCAACGCTGAGGCCCTGCGCGGGCGGGGGAAGATGGGTGGTGGGA 780
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 386 GlyAlaSerThrAlaThrAsnSerGluSerSerThrThrSerSerGlyValSerThrAla 405
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 781 AGGACATCTTCACCGCTCTCCATGATGAGAAGTTGACTTGGTGTATATAACAGAGATGG 840
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 406 ThrAsnSerGluSerSerThrThrSerSerGlyAlaSerThrAlaThrAsnSerAspSer 425
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 841 GAAAGTGTGCTCCTACGTTGACTTTTATTATATAACAGAAACACCAACCTGCTGCT 900
```

```
Db 426 SerThrThrSerSerGluAlaSerThrAlaThrAsn-SerGluSerSerThrValSerSe 445
Qy 901 CACAATACCCAGACAGCTCTACTAGCCAAAGTCTCTGTATGCAATTCATTATTCTCTGCTG 960
Db 445 rGlyIleSerThrValThrAsnSerGluSerSerThrThrSerSerGlyAlaAsnThrAl 465
Qy 961 GACAAACTCTGGAAGCAGCGTGTGAAACAGCGGGGGAAGAGAGCGGCTGAGCGGAG 1020
Db 465 aThrAsnSerGlySerSerVal-ThrSerAlaGlySerGlyThrAlaAlaLeuThrGlyM 485
Qy 1021 GGAGTGTGATGATTCCCAACCGACGAGCTGTCTCGCCTTTAGAACGTCGAGCCTCTCCA 1080
Db 485 etHisThr-----ThrSerHisSerAlaSerThrAlav 496
Qy 1081 TGCTGTATTACAAACAGCTCTCCACATTCGATTTCCATGCGCTGACCGTA---AGGATA 1137
Db 496 alSerGluAlaLysProGlySer-LeuValPro-----TrpGluIlePheLeuIle 513
Qy 1138 AAGCCTGTATATATGCAACTAGAAATGCTGCGCTTTTCAACCCCGTATTATTGATTTTA 1197
Db 514 ThrLeuValSerValValAlaAlaValGlyLeuPheAlaGlyLeuPheCysValarg 533
Qy 1198 TAGAGCTTTTCACTGGAAATCTACATAAATGTCAGTAAACCAATAAAGTTTCATTTCCTCA 1257
Db 534 AsnSerLeuSerLeuArgAsnThrPheAsnThrAlaValTyrHisProHisGlyLeuAsn 553
Qy 1258 AGGGGAATCAGGACGAGCCACACCCGCAATGGT 1290
Db 554 HisGlyLeu-----GlyProGlyProGlygly 562
```

RESULT 17

```
US-11-101-316-100
; Sequence 100, Application US/11101316
; Publication No. US20060099657A1
; GENERAL INFORMATION:
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: ANTIBODIES TO A POLYPEPTIDE ENCODED BY A NUCLEIC ACID
; FILE REFERENCE: P3230R1C17C1
; CURRENT APPLICATION NUMBER: US/11/101,316
; PRIOR FILING DATE: 2005-04-06
; PRIOR APPLICATION NUMBER: 10/063526
; PRIOR FILING DATE: 2002-05-03
; PRIOR APPLICATION NUMBER: 10/006867
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 09/380137
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: PCT/US99/12252
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 100
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-11-101-316-100
```

```
Alignment Scores:
  Pred. No.: 0.203      Length: 596
  Score: 101.50      Matches: 110
  Percent Similarity: 36.7%      Conservative: 49
  Best Local Similarity: 25.4%      Mismatches: 224
  Query Match: 2.3%      Indels: 50
  DB: 7      Gaps: 15
```

US-10-768-158-1 (1-2419) x US-11-101-316-100 (1-596)

Qy	61	AGTTTCGAGCAGCAAGTACT-----TCGAGTTCCATGGCGGTGGCGTGGCGCGC	105
Db	159	SerSerArgAlaSerThrAlaThrAsnSerGluSerSerThrLeuSerSerGlyAla---	177
Qy	106	CCTTCTGCCGGGAAGATGGAGGAGATCGCAACTCCCGGTGGCGGCCAGCAGCGT	165
Db	178	SerThrAlaThrAsnSerAspSerSerThrThrSerSerGlyAlaSerThrAlaThrAsn	197
Qy	166	GGATCGTCACCTACCCCAAGT-----CCGGCACACGCTTCTGCAGGAGGTGCTTACT	219
Db	198	SerGluSerSerThrThrSerSerGlyAlaSerThrAlaThrAsnSerGluSerSerThr	217
Qy	220	TGTTGAGCAGCGGGCTCACCCCATGAGATCGGCTTGATGAACATCCAGCAGCAGCTCC	279
Db	218	ValSerSerArgAlaSerThrAlaThrAsnSerGlu---SerSerThrThrSerSerGly	236
Qy	280	CGTCTCGAGTACCCACAGCGGGCTGGAGCATCATCAAGGAATGACCTCTCCCGGCC	333
Db	237	AlaSerThrAlaThrAsnSerGluSerArgThrThrSerAsnGlyAlaGlyThrAlaThr	256
Qy	340	TCATCAAGAGCC-----ACCTGCCCTACCCGCTTCTGCCCTCGACCTCCACAATGG	393
Db	257	AsnSerGluSerSerThrThrSerSerGlyAlaSerThrAlaThrAsnSerAspSerSer	276
Qy	394	ACTCCAAGGTCACTATATGCTCGCAACCCCAAGGATCTGGTGGTGCTTATTATCAGT	453
Db	277	ThrValSerSerGlyAlaSerThrAlaThrAsnSer-----GluSer	290
Qy	454	TCCACCGCTCTCTCGGGA---CCATGAGCTACCGAGGCACCTTTCAAGAATTTCCCGGA	510
Db	291	SerThrThrSerSerGlyAlaSerThrAlaThrAsnSerGluSerSerThrThrSerSer	310
Qy	511	GGTTTATGAATGATAAGCTGGGCTACGGCTCCTGGTTGACACGTCGACGAGTCTGGG	570
Db	311	Gly-----AlaSerThrAlaThrAsnSerAspSerSerSerThrThrSerSerGlyAla	327
Qy	571	AGCACCGCATGGATCGCAACGTGCTTTTCTCAAGTATGAAGACATGATCGGACCTGG	630
Db	328	GlyThrAlaThrAsnSerGluSer-----SerThrValSerSerGlyValSerThrVal	345
Qy	631	TGACGATGGTGGACGACGTGCCAGATTCTTGGGGGTGTCCTGTGACAAAGGCCACGTGG	690
Db	346	ThrAsnSerGluSerSerThrProSerSerGlyAlaAsnThrAlaThrAsnSerGluSer	365
Qy	691	AAGCCCTGACGGAGC-----ACTGCCAC-----AGCTGG	720
Db	366	SerThrThrSerSerGlyAlaAsnThrAlaThrAsnSerGluSerSerThrValSerSer	385
Qy	721	TGGACCAAGTGTGCAACGCTGAGGCCCTGCCCGTGGCGGGGAAGAGTTGGGTGTGGA	780
Db	386	GlyAlaSerThrAlaThrAsnSerGluSerSerThrThrSerSerGlyValSerThrAla	405
Qy	781	AGGACATCTTCCACGCTCCCATGAATGAGAAGTTTGACTTGGTGTATAAACAAGATGG	840
Db	406	ThrAsnSerGluSerSerThrThrSerSerGlyAlaSerThrAlaThrAsnSerAspSer	425
Qy	841	GAAGTGTGACCTCAGCTTGTGACTTTATTTATTAATACAGAAACAACAACCTGCATGCT	900
Db	426	SerThrThrSerSerGluAlaSerThrAlaThrAsn-SerGluSerSerThrValSerSe	445
Qy	901	CACAATACCCAGACAGTCTACTACCAAAAGTCTGTATGCAITTCATTTATCTTCTGCTG	960
Db	445	rGlyIleSerThrValThrAsnSerGluSerSerThrThrSerSerGlyAlaAsnThrAl	465
Qy	961	GACAACTCTGGAAGACGCTGTGAAACACGCGGGGAAGGAAAGCGGCGGTGACGGAG	1020
Db	465	aThrAsnSerGlySerSerVal-ThrSerAlaGlySerGlyThrAlaAlaLeuThrGlyM	485
Qy	1021	GGAGTGTGATGATTTCCCAACCGAAGCAGCTCTCTCGCCTTTAGAACGTGCAAGCTCTCCA	1080
Db	485	ethIleThr-----ThrSerHisSerAlaSerThrAlaVal	496

```

QY 1081 TGTCTGTAATACAAACAGTCTCCACATGTGACGTTCCTGAGCCGTA---AGGATA 1133
Db :||||: ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
496 a:SerGluAlaLysProGlySer-LeuValPro-----TrpGluIlePheLeuIle 513
QY 1138 AAGCCTGTAATATGCAACTAGAAATGCTGCGCTTTTCAACCCGTAATTATTGTTATTTA 1197
Db :||||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
514 ThrLeuValSerValValAlaAlaValGlyLeuPheAlaGlyLeuPhePheCysValArg 533
QY 1198 TAGAGCTTTTCACTGGAATCTACATAAATGTGAGTAAACCAATAAAGTTCTTCCTCA 1257
Db :||||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
534 AsnSerLeuSerLeuArgAsnThrPheAsnThrAlaValTyrHisProHisGlyLeuAsn 553
QY 1258 AGGGGAATCAGGAGCCACACCCCGAATGGT 1290
Db :||||: ||||| ||||| ||||| ||||| ||||| ||||| |||||
554 HisGlyLeu-----GlyProGlyProGlyGly 562

RESULT 18
US-10-509-131-35
; Sequence 35, Application US/10509131
; Publication No. US20060089493A1
; GENERAL INFORMATION:
; APPLICANT: Biogen Idec Inc.
; APPLICANT: Gately, Dennis
; TITLE OF INVENTION: NOVEL GENE TARGETS AND LIGANDS THAT BIND THERETO FOR TREATMENT OF COLON CARCINOMAS
; FILE OF INVENTION: DIAGNOSIS OF COLON CARCINOMAS
; FILE REFERENCE: 037003-0302886
; CURRENT APPLICATION NUMBER: US/10/509,131
; CURRENT FILING DATE: 2004-09-28
; PRIOR APPLICATION NUMBER: US 60/367,727
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 60/381,328
; PRIOR FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: US 60/386,747
; PRIOR FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US 60/427,564
; PRIOR FILING DATE: 2002-11-20
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35
; LENGTH: 836
; TYPE: PRN
; ORGANISM: Homo sapiens
US-10-509-131-35

Alignment Scores:
Pred. No.: 0.255 Length: 836
Score: 101.00 Matches: 166
Percent Similarity: 29.0% Conservative: 66
Best Local Similarity: 20.8% Mismatches: 264
Query Match: 2.3% Indels: 304
DB: Gaps: 39

US-10-768-158-1 (1-2419) x US-10-509-131-35 (1-836)
QY 2345 AAGATGCAAGCAGCAGACAGAGATCATGTGTAGTCTAGTAACATCCATACA----- 2292
Db :||||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
170 LysSerLysGlyArgArgGlyGlySerCysGlyAlaLeuAspThrLeuSerSerSer 189
QY 2291 -----GTACTAAAAATAGAAAAATATAAATGAAATTTTCAGAAAGCAGCCTC 2244
Db :||||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
190 ThrSerAspCysAlaIleCysLeuGluLysTyr----- 200
QY 2243 CCTCAGAAACACAGCAAGGTGCTCGCAGCCCAACCATGCGATGCCAGTCCAC--- 2187
Db :||||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
201 -----IleAspGlyGlyGluLeuArgValIleProCysThrHisArgPheHisArg 217
QY 2186 -----CTGCTGGCCGCTCTCCCATTTACAAACAAAAACAC 2148
Db :||||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
218 LysCysValAspProTrpLeuLeuGlnHisThrCysProHisCysArgHis----- 235
QY 2147 AAGCTTCTCTACACCCGTTTCAAAATACAGCACCAAGAACGAGAGCCCTTCCACAGCA 2088

```



```
QY 708 -----AGTGCTCCG-----TCAGGGCTTCAGCTGGCGCTTGTCACAGGACA 667
Db 149 ArgGluIleSerAlaSerLeuLeuTyrGlnAlaLeuProSerSerProAspHisGluSer 168
QY 666 CCCCAGGAATCTGGCCAGCTGCTCCACCATCGTCACAGGTCCCGATGCTTCAT 607
Db 169 ProSerGlnGluSerProAsnAlaProThrSerThr----- 180
QY 606 ACTTGAGAAAAAGCAGCTTCGAGTCCATCGCGTCTCCAGAACTCTCTGCACGTGCTCAA 547
Db 181 -----AlaValLeuGlySerTrpGlySerProGlnProSerLeuAlaPro 196
QY 546 ACCAGGACCTAGCCAGCTATCAT-----TCATAAACCTCCGGC----- 505
Db 197 ArgGluGlnGluAlaProGlyThrGlnTrpProLeuAspGluThrSerGlyIleTyr 216
QY 504 --AGAAATCTTGAAGGTGCTCGGTAGCTCA-----TGCTCCGCA 466
Db 217 ThrGluIleArgGluArgGluArgGluSerGluLysGlyArgProProTrpAlaAla 236
QY 465 GAGAGCGGTGAACATAATAAGACACACACAGATCCT---TGGGGTTGCGAGCCATAT 409
Db 237 ValValGly-----ThrProGlnAlaHisThrSerLeuGlnProHis 251
QY 408 AGATGACCTTGAGTCTCATTTG----- 384
Db 252 HisHisProTrpGlu-ProSerValArgGluSerLeuCysSerThrTrpProTrpLysAs 271
QY 383 -GAGTCAGAGGGCAGAAAGCGGTAGGCGAGGTCTTCATGAGCGGGAGAGGTCA 325
Db 271 nGluAspPheAsnGlnLysPheThr---GlnLeuLeuLeuGlnArgProHisProAr 290
QY 324 GTTCTTGATGATGCTCCAGCGCGGCTGTG-----GTACTCCAGGACCGGAGCT 274
Db 290 gSerGlnAspProLeuValLysArgSerTrpProAspTyrValGluGluAsnArgGlyHi 310
QY 273 GCTCGTCGATGTTCAAGCCGATCATCGGGTCAGCGCCCTGCTCACCAGTAGA 214
Db 310 sLeuIleGluIleArg-----AspLeuPheGlyProGlyLeuAspThrGlnGluProAr 328
QY 213 CCACCTCTCGCAGCAAGCTGTGCGGACTTGGG 180
Db 328 gIleValIleLeuGlnGlyAlaAlaGlyIleGly 339

RESULT 22
US-10-511-937-2934
; Sequence 2934, Application US/10511937
; Publication No. US2006008836A1
; GENERAL INFORMATION:
; APPLICANT: EXPRESSION DIAGNOSTICS, INC.
; APPLICANT: Wohlgemuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; APPLICANT: Prentice, James
; APPLICANT: Morris, MacDonald
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND MONITORING TRANSPLANT REJECTION
; FILE REFERENCE: 506612000104
; CURRENT APPLICATION NUMBER: US/10/511.937
; CURRENT FILING DATE: 2004-10-19
; PRIOR APPLICATION NUMBER: PCT/US2003/012946
; PRIOR FILING DATE: 2003-04-24
; PRIOR APPLICATION NUMBER: US 10/131,831
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US 10/325,899
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 3117
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2934
; LENGTH: 1429
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
US-10-511-937-2934
Alignment Scores:
Pred. No.: 0.617 Length: 1429
Score: 97.50 Matches: 70
Percent Similarity: 35.7% Conservative: 27
Best Local Similarity: 25.7% Mismatches: 105
Query Match: 2.2% Indels: 70
DB: 6 Gaps: 15

US-10-768-158-1 (1-2419) x US-10-511-937-2934 (1-1429)
QY 852 GGTCAACATTTCCCATCTTCTGTTTATACACCAAGTCAAACTTCTCATTCATGGAGACGG 793
Db 91 GlyHisSer---ProSerPheProTyrSerProSerGlu---ProHisLeuGlySerPro 108
QY 792 TGAAGATGTCCTTCCACAGCCCAACTCTTCCCGGCCACCGGAGGCGCTCAGGCTTGC 733
Db 109 SerGlnProThrSerThrAlaValLeuMetProTrpIleHisGluLeuProAlaGlyCys 128
QY 732 AGCACTGCT-----CCACCAGCTGGTGGC----- 709
Db 129 ThrGlnGlySerGluArgValLeuArgGlnLeuProAspThrSerGlyArgArgTrp 148
QY 708 -----AGTGCTCCG-----TCAGGGCTTCCAGCTGGCGCTTGTCACAGGACA 667
Db 149 ArgGluIleSerAlaSerLeuLeuTyrGlnAlaLeuProSerSerProAspHisGluSer 168
QY 666 CCCCAGGAATCTGGCCAGCTGCTCCACCATCGTCACAGGTCCCGATGCTTCAT 607
Db 169 ProSerGlnGluSerProAsnAlaProThrSerThr----- 180
QY 606 ACTTGAGAAAAAGCAGCTTCGAGTCCATCGCGTCTCCAGAACTCTCTGCACGTGCTCAA 547
Db 181 -----AlaValLeuGlySerTrpGlySerProGlnProSerLeuAlaPro 196
QY 546 ACCAGGACCGTACCCAGCTTATCAT-----TCATAAACCTCCGGC----- 505
Db 197 ArgGluGlnGluAlaProGlyThrGlnTrpProLeuAspGluThrSerGlyIleTyr 216
QY 504 --AGAAATCTTGAAGGTGCTCGGTAGCTCA-----TGCTCCGCA 466
Db 217 ThrGluIleArgGluArgGluArgGluSerGluLysGlyArgProProTrpAlaAla 236
QY 465 GAGAGCGGTGAACATAATAAGACACACACAGATCCT---TGGGGTTGCGAGCCATAT 409
Db 237 ValValGly-----ThrProGlnAlaHisThrSerLeuGlnProHis 251
QY 408 AGATGACCTTGAGTCTCATTTG----- 384
Db 252 HisHisProTrpGlu-ProSerValArgGluSerLeuCysSerThrTrpProTrpLysAs 271
QY 383 -GAGTCAGAGGGCAGAAAGCGGTAGGCGAGGTCTTCATGAGCGGGAGAGGTCA 325
Db 271 nGluAspPheAsnGlnLysPheThr---GlnLeuLeuLeuGlnArgProHisProAr 290
QY 324 GTTCTTGATGATGCTCCAGCGCGGCTGTG-----GTACTCCAGGACCGGAGCT 274
Db 290 gSerGlnAspProLeuValLysArgSerTrpProAspTyrValGluGluAsnArgGlyHi 310
QY 273 GCTCGTCGATGTTCAAGCCGATCATCGGGGTGAGCGCCCTGCTCACCAGTAGA 214
Db 310 sLeuIleGluIleArg-----AspLeuPheGlyProGlyLeuAspThrGlnGluProAr 328
QY 213 CCACCTCTCGCAGCAAGCTGTGCGGACTTGGG 180
Db 328 gIleValIleLeuGlnGlyAlaAlaGlyIleGly 339

RESULT 23
US-11-238-282-19
; Sequence 19, Application US/11238282
; Publication No. US20060089303A1
```


GENERAL INFORMATION:

; APPLICANT: Old, Lloyd J.
 ; APPLICANT: Scanlan, Matthew J.
 ; APPLICANT: Chen, Yao-Tseng
 ; TITLE OF INVENTION: Cancer-Testis Antigens
 ; FILE REFERENCE: L0461.70125US01
 ; CURRENT APPLICATION NUMBER: US/11/238,282
 ; CURRENT FILING DATE: 2005-09-29
 ; PRIOR APPLICATION NUMBER: US 10/054,683
 ; PRIOR FILING DATE: 2002-01-22
 ; PRIOR APPLICATION NUMBER: US 60/280,718
 ; PRIOR FILING DATE: 2001-03-30
 ; PRIOR APPLICATION NUMBER: US 60/285,154
 ; PRIOR FILING DATE: 2001-04-20
 ; PRIOR APPLICATION NUMBER: US 60/327,432
 ; PRIOR FILING DATE: 2001-10-05
 ; NUMBER OF SEQ ID NOS: 37
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 19
 ; LENGTH: 734
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-11-238-282-19

Alignment Scores:

Pred. No.: 0.576 Length: 734
 Score: 96.50 Matches: 179
 Percent Similarity: 30.8% Conservative: 93
 Best Local Similarity: 20.3% Mismatches: 282
 Query Match: 2.2% Indels: 329
 DB: 7 Gaps: 51

US-10-768-158-1 (1-2419) x US-11-238-282-19 (1-734)

```

Qy 597 -----TTTCTCAAGTATGAAGACATGCATCGGAGCTGGTG 632
Db 167 GlnSerAlaGluProGlnGlnAspPheAlaValIleGluMetHis-----Val 183
Qy 633 ACATGCTGGAGCAGCTGGCCAGATTCTCGGGGTCTCTGTGACAAAGGCCAGCTGGAA 692
Db 184 IleValGluLysGlnLeuTyrAsnHisMetGly-----SerAspThrThr 198
Qy 693 GCCCTGACGGAGCACTGCCACCCAGCTGGTGACCAAGTGTGCAACGCTGAGGCCCTGCC 752
Db 199 ValValAlaGlnLysValPheGlnLeu----- 207
Qy 753 GTGGCCGGGAAGAGTTGGCTGTGGAAGACATCTTCAACCTCTCATGAATGAGAAG 812
Db 208 -----IleGlyLeuThrAsnAlaIlePhe-----ValSerPheAsnIleThr 221
Qy 813 TTTGACTTG-----GTGTATAAACAAGAGATG-----GGAAG 845
Db 222 IleIleLeuSerSerLeuLeuLeuTrpIleAspGluAsnLysIleAlaThrThrGlyGlu 241
Qy 846 TGTGACCTCAGCTTTGACTTTTATTATATAATAACAGAAACAACACCTGCATGCTCAAA 905
Db 242 AlaAsnGluLeuLeuHisThrPheLeu-----Arg 251
Qy 906 TACCACAGACGTCTACTAGCCAAAAGTCTGTATGCATTATTATTTCTTCTGCGCAA 965
Db 252 TrpLysThrSerTyrLeuValLeuArgProHisAspValAlaPheLeuLeuValTyrArg 271
Qy 966 ACTCTGGAAGCAGCGTGTGAAACACGCGGGAAGGGAAG-----AGCGCGTGAGC 1016
Db 272 GluLysSerAsnTyrValGlyAlaThrPheGlnGlyLysMetCysAspAlaAsnTyrAla 291
Qy 1017 GGAGGGAGTGTGATGATTCACCAACCGAAGCAGC-----TGTCCTGCTTTT 1061
Db 292 GlyGlyValValLeuHisProArgThrIleSerLeuGluSerLeuAlaValIleLeuAla 311
Qy 1062 AGAAGCTGCAGCTCTCCATGCTCT--GATTACAAACAGTCTCCACATCCAGTTCCTCAAT 1118
Db 312 GlnLeuLeuSerLeuSerMetGlyIleThrTyrAspAspIleAsnLysCysGlnCysSer 331
Qy 1119 GCGCTGACCGCTAAGGATAAAGCCTGTAATATATATGCAACTAGAAATGTCTGCTTTTCAAC 1178
Db 332 GlyAla-----ValCysIleMetAsn 338
Qy 1179 CCGCTATTATGATTTT-----ATAGAGCTTTTCTACTGGA-----AAT 1217
Db 339 ProGluAlaIleHisPheSerGlyValLysIlePheSerAsnCysSerPheGluAspPhe 358
Qy 1218 CTACATAAATGCTAGTAACCAATAAAGTTTCATTTCCTCAAGGGAATTCAGAGCGAGCC 1277
Db 359 AlaHis-----PheIleSerLysGlnLysSerGlnCysLeu 370
Qy 1278 ACACCCGAATGGTAGAAAGATCTCAGGGTTAACTCTTTATTATTTTGTAGTTTATTATCTA 1337
Db 371 -HisAsnGlnProArg-----LeuAspProPhePheLysGlnAlaValCys 386
Qy 1338 AGGCACAGCCATTCTGTTCTCATTGTTCTGAGATAGTGTGAGAACAGAGGATGAGTT 1397
Db 386 sGlyAsnAlaLysLeu-----GluAlaGlyGluGluCysAspCysGlyThrGluGlnAspCys 405
Qy 1398 GGGTCTGTTGGGGG-----AATCTGCACACTGTGTTTATTCTGACCGAGTTCCTACT 1447
Db 405 sAlaLeuIleGlyGluThrCysCysAspIleAlaThrCysArgPheLysAlaGlySerAs 425
Qy 1448 TCTTCAGAACCTCTCTGAAATGAGCAGAAATTTGTTCTACTAGTCTTTCAGAAATGACGTC 1507
Db 425 nCysAlaGluGlyPro-----CysCysGluAsnCys-----LeuPheMetSerLysGluArgMe 443
Qy 1508 TTCTGCCAGAGACTTCCAGCGGGGCTCCAAAGGCCCAATGCGAGAGCGCCGCGGAGC 1567
Db 443 tCysArgProSerPheGluGluCysAspLeuProGluTyrCysAsnGlySerSerAlaSe 463

```

```
Qy 1568 ATGCTGCTGAGGAAGTCTG-----CCTGGTGAGGCTGGCAGGTGGAGTC 1612
Db 463 rCyProGluAenHisTyValGlnThrGlyHisProCyGlyLeuAenGlnTrp----- 481
Qy 1613 TAATGTCAGTCAGGAGCATTTTCATGTCAGTGGTGGAGAGTCGGCCACCAAGACCGAGT 1672
Db 482 -----lleCyHisle-AspGly----- 486
Qy 1673 TGCCTCGGAATTTGAGCTGAATTCACAGCCTTACTTTGTTCTGTAAGTATAGCCTA 1732
Db 487 -----ValCyMet----- 489
Qy 1733 CTAATGCTGGCAAGCAGATGCTTAATAGTAATTTCTAAATCCCGGGCTTTTATCATTT 1792
Db 490 -----SerGlyAspLysGlnCysThrAspThrPheGlyLysGluValGluPheGlyProS 508
Qy 1793 CAGTTGTTCTGTCGACCTCAGG--CGCTCAGCGCTGGGAGGA----- 1833
Db 508 erGluCySerHisLeuAenSerLysThrAspValSerGlyAsnCyGlyleSera 528
Qy 1834 -----CAATTTTGGAGTGTAGCCCTGTTCTACCTGGATCAGGT-----T 1873
Db 528 spSerGlyTyThrGlnCyGlyAlaAspAenLeuGlnCyGlyLysLeulleCysLysT 548
Qy 1874 GGCAGCGCCGCTGGTGTCTGTCCACTATCCTCGTGTATCTGAGGAGTAAAGGT 1933
Db 548 yrValGlyLysPheLeuLeuGlnleProArgAlaThrIleTyAlaAsnleSer- 567
Qy 1934 GAGTCTTTATTTGCTTCACTGCCCTAAATTTCT-----CACCCATTTCTGCTGAAGCAT 1987
Db 568 --GlyHisLeuCySilleAlaValGluPheAlaSerAspHisAlaAspSerGlnLysMetT 587
Qy 1988 GGAGAGTCGGGGGCGAGTAGCAGCAACCCCGTGGGACGGGGTGTCTGTCTCATTTAT 2047
Db 587 rpIleLysAspGly-----ThrSerCysGlySerAsnLysValCys----- 600
Qy 2048 GTGCTGGAAGACCAACCAAGTGTGTGTCAGGAGGTGCTGCTGTGAAGGGTCTCCG 2107
Db 601 -----ArgAsnGlnArgCysValSerSer---SerT 610
Qy 2108 TTCTTGTGTGTATTTGMAACGGGTGTA-----GAGAGAAGCTTGTGT----- 2151
Db 610 yrLeuGlyTyAspCysThrThrAspLysCysAsnAspArgGlyValCyAsnAenLysL 630
Qy 2152 -----TTTTTGTGTATTTGGGAGAGCGT 2176
Db 630 ysHisCysHisCysSerAlaSerTyLeuProProAspCysSerValGlnSerAspLeut 650
Qy 2177 GGCAGGAGGTGGCAGTCGTCATCGCATCGCATGTGGGCTC-----GGCAGCACCTTGGCTG 2230
Db 650 rpProGly-----GlySerIleAspSerGlyAsnPheProV 663
Qy 2231 TGTTTCTGTGAGGAGGCTGCTTCTGTGAAATTTCAATTTAT----- 2272
Db 663 alAlaIleProAlaArgLeuProGluArgG-TyrIleGluAenIleTyHisSerLys 682
Qy 2273 -----ATTTTTCTATTTTGTACTGTATGTGATGTACTGAGCTACAC 2317
Db 683 ProMetArgTrpProPhePheLeuPheIleProPhePhe----- 695
Qy 2318 ATGATCTCTGTGTC-----TTGCTGTGATCTTTTAATAAGACATGTTCCCGCGGTGCAAA 2374
Db 696 ---IleIlePheCysValLeulleAlaIleMetValLysValAenPheGlnArgLysLys 714
RESULT 24
US-10-505-928-50
; Sequence 50, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505.928
```

```
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 50
; LENGTH: 760
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: KIAA0582
US-10-505-928-50

Alignment Scores:
Pred. No.: 0.584 Length: 760
Score: 96.50 Matches: 143
Percent Similarity: 32.5% Conservative: 60
Best Local Similarity: 22.9% Mismatches: 228
Query Match: 2.2% Indels: 194
DB: 6 Gaps: 30

US-10-768-158-1 (1-2419) x US-10-505-928-50 (1-760)
Qy 1702 CTGTGGAATTCAGCTCAATTCGAGCGCAACTCGTCTTTGGTGCCGACTCTCCACC 1643
Db 58 ValTrpGlyAlaGluGlyIleProAlaProThrCysTrpIleGly----- 72
Qy 1642 CACTGCATGCAAAATGCTCTCGACTGCATTAGACTCCCACTGCCAGCCTCACCAGGCAGA 1583
Db 73 -----ThrAspProGlyGly 77
Qy 1582 CTTCCTCTCAGACATGTCGCGGGGT-----CCTCTGCATTTGGGCC 1541
Db 78 ProSerArgAlaHisGlnProGlnAlaSerAspAlaAsnArgGluProVal-AlaGluAr 97
Qy 1540 TTTGGAGCGCCGCTGGAAGTCTCTGGCAGAGGAGCTCCATTCTGAAGACCTAGTGAA 1481
Db 97 gSerGluProAlaLeuSerGlyLeuProProAlaThrMetGlySerGlyAspLeuLeu-- 116
Qy 1480 CAATTTCTGCTCATTTCTGAGGAAGTCTGGAAGTGAAGTCTCCGTCAGATAAAACAAGTG 1421
Db 117 -----LeuSerGlyGluSerGlnValGluLysLeuSerSerSe 131
Qy 1420 TCCAGATTTCCCAACAGACACCAACTC-----ATCCTCTGTTCTCACCCTA 1373
Db 131 rGluGluPheProGlnThrLeuSerLeuProArgThrThrIleCysSerGlyHisAs 151
Qy 1372 TCTCAGAACCAAGTGAACAGATGGCTGT-----GCCTTAGATAATAAAC 1325
Db 151 pAlaAspThrGluAspAspProSerLeuAlaAspLeuProGlnAlaLeuAspLeuSerGl 171
Qy 1324 TACAAAAATAAAGAGTTAAACCTGAGATCTTTTCTACCATCTGGGTGGCTCGCTCTGA 1265
Db 171 nGlnProHisSerSer-----GlyLeuSerCysLeuSerGlnTrpLysSerValLe 188
Qy 1264 TTCCCTTGGAAATGAACATTTTATTTGGTTTACTGACATTTATAGATTTCAGTGAAA 1205
Db 188 uSerPro----- 190
Qy 1204 AGCTCTATAAATACATAATACGGGGTTGAAAAGGCAGACATTTCTAGTTGCATATATTA 1145
Db 191 -----GlySerAlaAlaGlnProSerSerCys-SerIle- 201
Qy 1144 CAGGCTTTATCTTACCGTCCAGGCCATTCGAATGCAATGTGGAGACTGTTTCTAATCA 1085
Db 202 -----SerAlaSerSerThrGlySerSerLeuGln-----G 212
Qy 1084 GACATGGAGAGGCTGCACGTTTAAAGCGCAGACAGCTGCTTCGGTTGGGAATCATCACA 1035
Db 212 lyHisGlnGluArgAlaGluProArgGlyGlySerLeuAlaLysVal---SerSer-Ser 230
Qy 1024 CTCCTCCGCTCAGCGCGCTCTTCCCTTCCCGCGCTGTTTCA-----CACGCTGCT 974
```

```
Db      231 LeuGluProValValProGlnGluProSerSerValValGlyLeuGlyProArgProGln 250
Qy      973 TCCAGAGTTGTCACAGGAAGAAATAATGAATGATACAGGACTTTTGGCTAGTAGACTG 914
Db      251 TrpSerProGlnProValPheSerGlyGlyAspAlaSerGlyLeuGlyArgArgLeu 270
Qy      913 TCTGGGTATTGTGAGCATGCAGGTGTGTCTTCTGTTATTATAATAAAAGTCAACGCTG 854
Db      271 SerPheGlnAlaGluTyrTrpAlaCys-Val- 280
Qy      853 AGGTACACACTTCCCATCTCTGTTTATACACCAAGTCAAACTTCTCATTTCA-----TG 800
Db      281 -----LeuProAspSerLeu---ProProSerProAspArgHisSerProLeuTr 296
Qy      799 GAGACGTTGAAGATGCTCTTCC-----ACAGCCCAACT 767
Db      296 pAsn-----ProAsnLysGluTyrGluAspLeuLeuAspYrThrTyrProLe 312
Qy      766 CTTCCCGCGCCACGGGCGGCTCAGCGTTGCAGCACTGGTCCACGAGCT-----715
Db      312 uArgProGlyProGlnLeu---ProLysHisLeuAspSerArgValProAlaAspProVa 331
Qy      714 -----GGTGCAGTGTCCGTCAGGCTTCCAGCTGGGCTTGTTCAC-----AG 671
Db      331 lLeuGlnAspSerGlyValAspLeuAspSerPheSerValSerProAlaSerThrLeuLy 351
Qy      670 GACACCCCGGAGATCTGGCCAGCTGTCTCCACCATCGTCACAGGTCCTCGATGATGTCT 611
Db      351 sSerProThrAsnValSerProAsnCysProProAla-----363
Qy      610 TCATACTTGAGAAAAGACGCTTCAGATCCAGTCCGCTCCAGAACTCCGTGCAGCGTC 551
Db      364 -----GluAlaThrAlaLeuProPheSerGlyProArgGluProSerLeuLy 379
Qy      550 TCAAAACAGGAGCGGTAGCCAGCTTATCATTTATAAACCTCCGCGCAATCTTTGAAAG 491
Db      379 sGlnTrpProSerArgValProGlnLysGln-----GlyGlyMetGlyLeu---- 394
Qy      490 GTGCTTCGGTAGCTATGGTCGCGAGAGAGCGGTGGAACTGATATAAGACACACCCAGA 431
Db      395 -----AlaSerTrpSerGlnLeuAla-----SerThrPro-----404
Qy      430 TCCTTGGGGTTCGAGCCATATAGATGACCTTGGAGTCTCCATCTGGAGGTACAGGGC 371
Db      405 -----ArgAlaProGlySerArgAspAlaArgTrpGluArgAr 417
Qy      370 AGAAAGCGGTAGGCGAGGTGGCTCTTGATGAGCGGGGAGAGGTCACTTCTTGATGATG 311
Db      417 gGluProAlaLeuArgGly-AlaLysAspArgLeuThrIleGlyLysHisLeuAsp-Met 436
Qy      310 TCCAGGCCCGCTGTGGTACTCCAGGACCGGAGCTGCTCGTGTGATGTTTCATCAAGCG 251
Db      437 Gly-SerProGlnLeuArgThrArgAspArgGly-----TrpProSerProAr 452
Qy      250 ATCTCATCGGGTCAGCGCCCTGGCTCACCAGTAGACACCTCTCTGAGCAAGCTGTG 191
Db      452 gProGluArgGluLysArgThrSerGlnSerAlaArgProThrCysThrGlu-----470
Qy      190 CCGGACTTGGGGTA-----GGTGAGCATCCACACGTCGCTGGGCGCG 149
Db      471 -SerArgTrp-LysSerGluGluGluValGluSerAspAspGluTrLeuAlaLeuProA 490
Qy      148 ACCGG-----GAAGTTGGCGATCTCTCC 125
Db      490 laArgLeuThrGlnValSerSerLeuValSerTyrLeuGlySerIleSerThrLeuVal 510
Qy      124 ATCTTCCCGCGCAGAA-----GGCGCGCAGCGCCACGCGCATGAACTCGAAGTACTG 71
Db      510 hrLeuProGlyAspIleLysGlyGlnSerPro-----LeuGluValSerA 526
Qy      70 CTCTCGAACTCCCGCGGGGTGTGGGGTCTCGGGCTCTCGGCTCGCTCTCCGCATCGCGCG 14
Db      526 sp-SerAspGlyProAlaSerPheProSerSerSerSerGlnSerGlnProPro 544
```

RESULT 25

```
US-10-505-928-227
; Sequence 227, Application US/10505928
; Publication No. US2006008532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178 US/10/505,928
; CURRENT APPLICATION NUMBER: 2004-08-27
; PRIOR FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 227
; LENGTH: 1663
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-227

Alignment Scores:
Pred. No.: 0.793 Length: 1663
Score: 96.50 Matches: 100
Percent Similarity: 35.5% Conservative: 70
Best Local Similarity: 20.9% Mismatches: 174
Query Match: 2.2% Indels: 135
DB: 6 Gaps: 23

US-10-768-158-1 (1-2419) x US-10-505-928-227 (1-1663)
Qy      1318 AATAAGAGTTAAACCTGAGATCTTTTACCATTCGGGTGGTGGCTCGTCTGATTCCCC 1259
Db      329 AsnLysGluLeuPheLeuGlnSerHisThrGluIleGlyValSer-----343
Qy      1258 TTGGAATGAACATTTTATTGTTTACTGCATTTATGTAGATTTCAGTGGAAAAGCTCT 1199
Db      344 -----TyrGlnTyrAlaLeuAspLeuGln 351
Qy      1198 ATAATAATACATAATACGGGTTCGAAAGGAGACATCTTAGTTGCATATATTACAGCT 1139
Db      352 ThrGlnHisAsnHisPheAlaMetAsnSerMetAsnAlaTyrValAsnIleAsnArgIle 371
Qy      1138 TTATCCTTACGTCACGACCATTTGGAACGTG-----CAA 1106
Db      372 MetSerValAlaSerArgLeuSerGluAlaGlyHisTyrAlaSerGlnGlnIleLysGln 391
Qy      1105 TGTGAGACACTGTTTGTATATCAGACATCGAGAGGCTGCACGTTCTTAAAGCGGAGACAGCTG 1046
Db      392 -IleSerThrGlnLeuAspGlnGluTrpLys-----401
Qy      1045 CTTCCGTTGGAAATCATCACACTCCCTCCGTCACGGCTCTTCCCTTCCCGCTGTT 986
Db      402 ----SerPheAlaAlaLeuAspGluArgSerThrIleLeuAlaMetSerAlaValph 420
Qy      985 TCACACGCTGCTCCAGAGTTGTCCACAGGAATAAATGAATGCATACAGGACTTTTG 926
Db      420 eHisGlnLysAlaGluGlnPheLeu---SerGlyVal-----AspAlaTr 434
Qy      925 GCTAGTAGACTGCTG-----GGTATTGTGAGCATGCAGGTGTGTTCTTCTATTAT 872
Db      434 pCysLysMetCysSerGluGlyGlyLeuProSerGluMetGlnAspLeuGluLeuAla 454
Qy      871 AAATAAAGTCAACAGGTGAGGTACACTTTCCCATCTCTCTTTAT-----ACACC 821
Db      454 eHisHisGlnThr-----LeuTyrGluGlnValThrGl 466
Qy      820 AAGTCAAACTCTCATTCATCATGAGACGCTGAA-----GATGTCCTTCCACAGCCCA 770
Db      466 nAlaTyrThrGlu-ValSerGlnAspGlyLysAlaLeuLeuAspValLeuGlnArgPro- 485
Qy      769 ACTCTTCCCGCGCCACGGGCGGCGCTTCAGCGTTGCAGCA-----CTGGTCC 722
```

```
Db 486 --LeuSerProGlyAsnSerGluSerLeuThrAlaThrAlaAsnTyrSerLysAlaValH 505
Qy 721 ACCAGCTGGTGGCAGTGTCCGTCCAGGCTTCCAGCTGGCGCTTGTACAGGACACCC 662
Db 505 isGlnValLeuAspValValHis-----512
Qy 661 AGGAATCTGGCCAGCTGTCCACCATCTCCAGGTCCTCCAGTGCATGCTTTCATCTTG 602
Db 513 -----GluValLeuHisGlnArgLeuGluSer-----1letpG 526
Qy 601 AGAAAAAGCAGCTTCCAGTCCATGCTCCAGAACTCTCCAGCACTCTCCAGCTCAAAACCAG 542
Db 526 lnHisArgLysValArgLeuHisGlnArgLeuGlnLeuCysValPheGlnGlnAspValG 546
Qy 541 GAGCCGTAGCCAGCTTATCATTCATAAACCTCCGGCAGAAATTCCTGAAAGTGCCTCGG 482
Db 546 lnGlnValLeuAspTrpIleGluAsnHisGlyGluAlaPheLeuSerLysHisThrGlyV 566
Qy 481 TAGCTCATGTCGGCAGAGCGGTGGAACTGATTAAGACACACACAGATCCTTGGGG 422
Db 566 alGlyLysSerLeuHisArgAlaArgAlaLeuGlnLysArgHisAspPheGluGluV 586
Qy 421 TTGGGAGCATATAGATCACCTTGGAGTCTCCATTTGGAGTCCAGAGGCGAGAAAGCG- 363
Db 586 alAlaGlnAsnThrTyrThrAsnAlaAspLysLeuLeuGluAlaAlaGluGlnLeuAlaG 606
Qy 362 --GTAGGCGAGGTGGCTTGTATGATGAGCGGGGAGAGGT-----327
Db 606 lnThrGlyGluCysAspProGluGluIleTyrLysAlaAlaArgHisLeuGluValArgI 626
Qy 326 ----CAGTTCTTGATGATGTCACGGCCCG-----300
Db 626 leGlnAspPheValArgValGluGlnArgLysLeuLeuLeuAspMetSerValSerP 646
Qy 299 -----CTGTGGGTA---CTCCAGGAC-----CGGGAGCTGCTCG 269
Db 646 heHisThrHisThrLysGluLeuThrTrpMetGluAspLeuGlnLysGluMetLeuG 666
Qy 268 TCGATGTTTCATCAAGCCGATCTATCGGGGTGAGCGCCCTGGCTCACCAAGTAGACACC 209
Db 666 luAspVal---CysAlaAsp-----SerValAspAlaValGlnGluLeuLysGlnP 683
Qy 208 TCCTGCAGACAGCTGTGCGGACTTGGGTA-----GGTAGC 170
Db 683 heGlnGlnGlnThrAlaThrLeuAspAlaThrLeuAsnValIleLysGluGlyGluA 703
Qy 169 AT-----CCACAGCTGCTGGCGCCGACCGGGAAGTGGCG 134
Db 703 spLeuIleGlnLeuArgSerAlaProProSerLeuGlyGluProSerGluAlaArgA 723
Qy 133 ATCTCTCTCATCTTCCCGCGCAGAAAGGCGGCGAGCCGACGCGCATGGAACTC 81
Db 723 sp-----SerAlaValSerAsnAsnLysThrProHisSerSerSerile 737

RESULT 26
US-11-140-487A-772
; Sequence 772, Application US/11140487A
; Publication No. US20060093617A1
; GENERAL INFORMATION:
; APPLICANT: Imogenetics N.V.
; TITLE OF INVENTION: Peptides for inducing a CTL and/or HTL response to Hepatitis C vi
; FILE REFERENCE: 166
; CURRENT FILING DATE: 2005-05-31
; PRIOR APPLICATION NUMBER: US/11/140,487A
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: EP 04012951.2
; PRIOR FILING DATE: 2004-10-28
; PRIOR APPLICATION NUMBER: EP 04447239.7
; PRIOR FILING DATE: 2005-03-25
; PRIOR APPLICATION NUMBER: EP 05102441.2
; PRIOR FILING DATE: 2005-03-25
; PRIOR APPLICATION NUMBER: US 60/576,310
; PRIOR FILING DATE: 2004-06-03
; PRIOR APPLICATION NUMBER: US 60/622,782
```

```
; PRIOR FILING DATE: 2004-10-29
; PRIOR APPLICATION NUMBER: US 60/665,395
; PRIOR FILING DATE: 2005-03-25
; NUMBER OF SEQ ID NOS: 2278
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 772
; LENGTH: 3020
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-11-140-487A-772

Alignment Scores:
Pred. No.: 1 Length: 3020
Score: 96.50 Matches: 118
Percent Similarity: 31.7% Conservative: 48
Best Local Similarity: 22.5% Mismatches: 156
Query Match: 2.2% Indels: 203
DB: 7 Gaps: 28

US-10-768-158-1 (1-2419) x US-11-140-487A-772 (1-3020)
Qy 1175 GAAAGGCGAGACATCTAGTTGTCATATATACAGGCTTTATCTTACGTCAGGCCATT 1116
Db 1983 AspMetValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuValValGlyValIle 1902
Qy 1115 GGAACCTCAATGTGGAGACTGTTTGTAAATCAGACAT-----GGAGAGGCTGCA--- 1068
Db 1903 CysAlaAlaIleLeuArg-----ArgHisValGlyProGlyGluGlyAlaVal 1918
Qy 1067 -----CGTCTAAAGCGGAGACAGCTGCTTCGTTGGGAATCATCACATCCCTCGCT 1014
Db 1919 GlnTrpMetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHis 1938
Qy 1013 CAC-----GCCGCTCTTCCCTCCCGCTGTTTC 984
Db 1939 TyrValProGluSerAspAlaAlaArgValThrAlaLeu--LeuSerSerLeuThrV 1958
Qy 983 ACAGCTGCTTCCAGAGTTTGTCCAGCAAGAAATAATGAA-----TGCA 939
Db 1958 alThrSerLeuLeuArgLeuHisGlnTrpIleAsnGluAspTyrProSerProCys 1978
Qy 938 TACAGGACTTT-----928
Db 1978 erAspAspTrpLeuArgIleIleTrpAspTrpValCysSerValLeuThrAspPheLysT 1998
Qy 927 --TGCTA-----GTACACTGCTCG- 910
Db 1998 hrTrpLeuSerAlaLysIleMetProAlaLeuProGlyLeuProPheIleSerCysGlnL 2018
Qy 909 -----GGTATTGTGAGCATGCAGGTGTTGTTTCTG 879
Db 2018 ysGlyTyrLysGlyValTrpArgGlyAspGlyValMetSerThrArgCysProCysGlyA 2038
Qy 878 TTATTATAATAAAGTCAAAAGTGGAGTCA-----847
Db 2038 laThrIleThr-----GlyHisValLysAsnGlySerMetArgLeuAlaG 2053
Qy 846 -----ACTTTCCCATCTTCTGTTTATACACCAA 819
Db 2053 lyProArgThrCysAlaAsnMetTrpHisGlyThrPheProIleAsnGluTyrThrThrG 2073
Qy 818 GTCAAACTTCTCATTCATGAGACGGTGAAGATGCTTCCACAGCCCAACTCTTCCCG 759
Db 2073 ly-----ProSerThr-ProCysProSerPro 2081
Qy 758 GCCCAGCGGCGGCTC-----ACGGTTGACAGCACTGTGTCCACAGCTGGTGGCAGTGCTC 702
Db 2082 AsnTyrThrArgAlaLeuTrpArgValAlaAla-----AsnSerTyrValGluValArg 2099
Qy 701 CGTCAGGCTTCCAGCTGGGCTTGTCCACAGCACACCCCGAGGAATCTGGCCAGCTGCTC 642
Db 2100 ArgValGly---AspPheHisTyrIleThrGlyAlaThrGluAspGlu-----Leu 2115
```

```
Qy 641 CACCATCGTCACAGGTCCTCCGATCGATCTCTTATCTTGAGAAAAGCAGCGTTCAGTC 582
Db 2116 LysCysProCysGlnValProAlaGluPheThrGluAlaAspGlyValArgLeu 2135
Qy 581 CAT----- 579
Db 2136 HisArgTyrAlaProProCysLysProLeuLeuArgAspGluIleThrPheMetValGly 2155
Qy 578 -----GCGGTG 573
Db 2156 LeuAsnSerTyrLeuIleGlySerGlnLeuProCysGluProGluProAspValSerVal 2175
Qy 572 CTCCAGACTCCG-----CACGTGCTCAAAACAGGAGCGGTAGCCAGCTT 525
Db 2176 LeuThrSerMetLeuArgAspProSerHisIleThrAlaGluThrAlaAlaArgArgLeu 2195
Qy 524 ATCATTCATAAACCTCCGGCAGAGATTCTTGAAGGTGCTCGGTAGTCATGTGTCGCGAG 465
Db 2196 AlaArgGlySerProProSerGluAlaSerSerSerAlaSerGlnLeuSerAlaPro--- 2214
Qy 464 AGACGGTGGAACTGATAAAGACACACAGATCCTTGGGGTTGGAGCCATATAGAT 405
Db 2215 ---SerLeuLysAlaThrCysGlnThrHisArgProHisProAspAlaGluLeuValAsp 2233
Qy 404 -----GACCTGGAGTCTCCATTGTGGAGGTTCAGAGGGCAGAAA 366
Db 2234 AlaAsnLeuLeuTyrArgGlnGluMetGlySerAsnIleThrArgValGluSerGluThr 2253
Qy 365 GCGGTAGGCGAGGTGGCTTTGATGAGCGGGGAGAGGTTCAGTTCTTGTATGTATCCAG 306
Db 2254 LysVal-----ValIleLeuAspSer-----PheGluProLeuArg 2265
Qy 305 GCCGGCTGGGGTACTCCAGACGGGAGGTGCTGTCGA-----TGTTATCAAGC 253
Db 2266 AlaGlu-----ThrAspAspAlaGluLeuSerValAlaAlaGluCysPhe----- 2280
Qy 252 CGATCTCATCGGGTACGCGCTGGCTCACAAGTAGACACCTCTCGCAGCAAGCTGG 193
Db 2281 -----LysLysPro-----ProLysTyrProProAlaLeuProIleTyr 2293
Qy 192 TGCGGAGCTGGGGTAGTGCAGTCCACACGTCGCTGGGCGCGACCGGGAAGTTGGCGA 133
Db 2294 AlaArgProAspTyrAsnProProLeuLeuAspArgTyrLysAlaProAsp-----Tyr 2311
Qy 132 TCTCTCATCTCCCGGCGAGAGGGCGCAGCGCCACGCCATCGAAGTACT 73
Db 2312 ValProProThrValHisGly----- 2318
Qy 72 TGCTCTGMACTCCCGGGGTGCTGGGGGTCTCGGCTCGCTCTCGCCATGCGCGC 13
Db 2319 CysAlaLeuProProArgGlyAla-----ProProValProProProArgArg 2334
Qy 12 CGTCGC 7
Db 2335 LysArg 2336
RESULT 27
US-10-196-749-238
; Sequence 238, Application US/10196749
; Publication No. US20060094864A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Deanoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
```

```
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C340
; CURRENT APPLICATION NUMBER: US/10/196,749
; CURRENT FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 238
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-196-749-238
Alignment Scores:
Pred. No.: 0.461 Length: 325
Score: 96.00 Matches: 70
Percent Similarity: 29.5% Conservative: 20
Best Local Similarity: 23.0% Mismatches: 104
Query Match: 2.2% Indels: 111
DB: 6 Gaps: 15
US-10-768-158-1 (1-2419) x US-10-196-749-238 (1-325)
Qy 850 TCACACTTTCCCATCTCTCTTTATACACCAAGTCAACTTCTCATTCATGGAGACGGTG 791
Db 58 SerGluHisProGlnProAlaLeuAspProArgSerAsnAspLeuAlaArgValProLeu 77
Qy 790 AAGATGTCCTTCCACAGCCCAACTCTTCCCGGCCACGGGCGGCTCAGCGTTCGAG 731
Db 78 LysLeuSerVal---ProProSerAspGlyPheProProAlaGlyGlySerAlaValGln 96
Qy 730 CACTGGTCCACCGACTGG----- 710
Db 97 ArgTyrProProSerTyrGlyLeuProAlaMetAspSerTyrProProGluAspProTyr 116
Qy 709 CAGTGTCCGTCAGGGCTTCAGCTGGGCTGTGTCAGGACGACACCCCGAGAAAT----- 656
Db 117 GlnMetMetAlaAlaAlaAlaGluAspArgLeuGlyGluAlaLeuProGluGluLeuSer 136
Qy 655 ---CTGGCCAGCTGCTCCACCATCGTCACAGGTCCCGATGTCATGTCTTCTCATCTGAGA 599
Db 137 TyrLeuSerSerAlaAlaAlaLeu-AlaProGly----- 147
Qy 598 AAAAGCAGGTTCGAGTCCATGCGGTGCTCCAGAACTCTCCAGAACTCTCCAGCTGCTCAAACCCAGAG 539
Db 148 -----SerGlyProLeuProGlyGluSerSerProAspAlaThrGlyLeuSerPr 164
Qy 538 CCGTAGCCAGCTATTCATTATCAAACTCCGGCAGAAATCTTGAAGGTGCTCGGTAG 479
Db 164 oGluAlaSerLeuLeuHisGlnAspSerGluSerArgArgLeuProArg----- 180
Qy 478 CTCATGGTCCGACAGAGCGGTGAACTGTATAAAGACACCAACCATCTTGGGGTTG 419
Db 181 -SerAsnSerLeuGlyAlaGlyGlyLysIleLeuSerGlnArgPro----- 195
```

QY 418 CGAGCCATATAGATGACCTTGAGTCTC-----CATTGTGGAGG 380
Db 196 -----ProTrpSerLeuIleHisArgValLeuProAspHisProTrpG1 210
QY 379 TCAGAGGCGAAGCGGTAGGCGAGTGGCTCTTGATGAGCGGGAGAGGTCAAGTCC 320
Db 210 yThrLeuAsnProSerValSerTrpGlyGly----- 220
QY 319 TTGATGATGTCAGGCGCGGTGTGGTACTCCAGAGCGGAGGCTGCTCGTCGATGTC 260
Db 221 -----GlyGlyProGlyThrGly----- 226
QY 259 ATCAAGCCGATCTCATCGGGTCAAGCGCTGGGCTCAACAAGTAGACCACTCTCGCAGC 200
Db 227 -----TrpGlyThrArgPro-----MetProHisProGluG1 237
QY 199 AAGCTGG-----TGCGGACTTGGGTAGGTGACGATCCACAG 161
Db 237 yIleTrpGlyIleAsnAsnGlnProProGlyThrSerTrpGly-----AsnIleAs 254
QY 160 TCCTGTGGCGCACCGGGAAGTTGG-----CGATCTCTCCATCTTCCCGCGG 113
Db 254 nArgTyrProGly---GlySerTrpGlyAsnIleAsnArgTyrProGlyGlySerTrpG1 273
QY 112 CAGAAGGCGCGAGCGCACGCCATCGAACTCGAAGTACTTCTCTCGAACTCCCGCGG 53
Db 273 y-----AsnIleAsnArgTyrProGlyG1 281
QY 52 GTGCTGGGGTCT 40
Db 281 ySerTrpGlyAsn 285

RESULT 28

US-11-101-316-64
; Sequence 64, Application US/11101316
; Publication No. US20060099657A1
; GENERAL INFORMATION:
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: ANTIBODIES TO A POLYPEPTIDE ENCODED BY A NUCLEIC ACID
; TITLE OF INVENTION: UNDEREXPRESSED IN MELANOMA
; FILE REFERENCE: P3230R1C17C1
; CURRENT APPLICATION NUMBER: US/11/101,316
; CURRENT FILING DATE: 2005-04-06
; PRIOR APPLICATION NUMBER: 10/063526
; PRIOR FILING DATE: 2002-05-03
; PRIOR APPLICATION NUMBER: 10/006867
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 09/380137
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: PCT/US99/12252
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 64
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Homo Sapien
US-11-101-316-64

Alignment Scores:

Pred. No.: 0.461 Length: 325
Score: 96.00 Matches: 70
Percent Similarity: 29.5% Conservative: 20
Best Local Similarity: 23.0% Mismatches: 104
Query Match: 2.2% Indels: 111

Db: 7 Gaps: 15

US-10-768-158-1 (1-2419) x US-11-101-316-64 (1-325)
QY 850 TCACACTTCCCATCTCTTGTATTATACACCAAGTCAAACTTCTCATTTCATGAGACGGTG 791
Db 58 SerGluHisProGlnProAlaLeuAspProArgSerAsnAspLeuAlaArgValProLeu 77
QY 790 AAGATGTCCTTCCACAGCCCAACTCTTCCCGCGCCACAGGGCGCTCAGCGTTCAG 731
Db 78 LysLeuSerVal---ProProSerAspGlyPheProProAlaGlyGlySerAlaValGln 96
QY 730 CACTGTGTCACCACTGG-----TGG 710
Db 97 ArgTrpProProSerTrpGlyLeuProAlaMetAspSerTrpProProGluAspProTrp 116
QY 709 CAGTGCTCCGTCAGGGCTTCCAGCTGGGCTTGTACAGGACACCCCGAGGAAT----- 656
Db 117 GlnMetMetAlaAlaAlaGluAspArgLeuGlyGluAlaLeuProGluGluLeuSer 136
QY 655 ---CTGGCCAGCTCTCCACCATCGTCACCAAGTCTCCCGATCGATGCTCTTCATCTTGAGA 599
Db 137 TyrLeuSerSerAlaAlaLeu-AlaProGly----- 147
QY 598 AAAAGCAGTTCAGTCCATCGCTCCCGGTCTCCAGAACTCTCTGCACGTCTCAACACGAGAG 539
Db 148 -----SerGlyProLeuProGlyGluSerSerProAspAlaThrGlyLeuSerPr 164
QY 538 CCGTAGCCAGCTTATCATTAACCTCCGGCAGAACTTCTTGAAGGTCCCTCGGTAG 479
Db 164 oGluAlaSerLeuLeuHisGlnAspSerGluSerArgLeuProArg----- 180
QY 478 CTGATGTCGCGCAGAGCGGTGAACTGATAATAAGACACACACAGATCCTTGGGGTTG 419
Db 181 -SerAsnSerLeuGlyAlaGlyLysIleLeuSerGlnArgPro----- 195
QY 418 CGAGCCATATAGATGACCTTGGAGTCTC-----CATTGTGGAGG 380
Db 196 -----ProTrpSerLeuIleHisArgValLeuProAspHisProTrpG1 210
QY 379 TCAGAGGCGAAGCGGTAGGCGAGTGGCTCTTGATGAGCGGGAGAGGTCAAGTCC 320
Db 210 yThrLeuAsnProSerValSerTrpGlyGly----- 220
QY 319 TTGATGATGTCAGGCGCGGTGTGGTACTCCAGAGCGGAGAGTCTCTCGTCGATGTC 260
Db 221 -----GlyGlyProGlyThrGly----- 226
QY 259 ATCAAGCCGATCTCATCGGGTCAAGCGCTGGTCTCAACAAGTAGACCACTCTCGCAGC 200
Db 227 -----TrpGlyThrArgPro-----MetProHisProGluG1 237
QY 199 AAGCTGG-----TGCGGACTTGGGTAGGTGACGATCCACAG 161
Db 237 yIleTrpGlyIleAsnAsnGlnProProGlyThrSerTrpGly-----AsnIleAs 254
QY 160 TCCTGTGGCGCGCACCGGGAAGTTGG-----CGATCTCTCCATCTTCCCGCGG 113
Db 254 nArgTyrProGly---GlySerTrpGlyAsnIleAsnArgTyrProGlyGlySerTrpG1 273
QY 112 CAGAAGGCGCGCAGCGCACGCCATCGAAGTACTTGTCTCTCGAACTCCCGCGG 53
Db 273 y-----AsnIleAsnArgTyrProGlyG1 281
QY 52 GTGCTGGGGTCT 40
Db 281 ySerTrpGlyAsn 285

RESULT 29

US-10-199-229-13
; Sequence 13, Application US/10199229
; Publication No. US20060099701A1
; GENERAL INFORMATION:

```
; APPLICANT: She, Jin-Xiong
; APPLICANT: Kumar, Pradeep
; APPLICANT: Wang, Cong-Yi
; TITLE OF INVENTION: ACTIVATORS OF CYCLIN-DEPENDENT KINASES
; FILE REFERENCE: 5853-210
; CURRENT APPLICATION NUMBER: US/10/199,229
; CURRENT FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 951
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-199-229-13

Alignment Scores:
Pred. No.: 0.85 Length: 951
Score: 95.00 Matches: 69
Percent Similarity: 34.3% Conservative: 23
Best Local Similarity: 25.7% Mismatches: 95
Query Match: 2.2% Indels: 81
DB: 6 Gaps: 13

US-10-768-158-1 (1-2419) x US-10-199-229-13 (1-951)
Qy 783 CCTTCCAGCCCAACTCTCCCGCCGCGGCGGCGGCGCTCAGCGTTGTCAGCACTGGT 724
Db 81 ProValProAlaAlaProValProAlaProSerLeuAlaProGlyGluAenGlyThrGly 100
Qy 723 -----CCACCAGCTGTGGCGAGTCTCCG 700
Db 101 AspTrpAlaProArgLeuValPheIleGluProProGlyAlaGlyGlyAlaAlaPro 120
Qy 699 TCAGGGCTTCAGCTGGGCGCTTGTACAGGACACCCCGGAGAACTGCGCCAGCTGCTCCA 640
Db 121 SerAlaValProThrArgPro----- 127
Qy 639 CCATCGTCACGAGTCCGATGTCATCTTATCTTGAAGAAAGCAGCTTCGAGTCCA 580
Db 128 -----ProGlyProGlnArgCys-----ArgGluGlnSerAsp-Tr 139
Qy 579 TCGGTGTCTCCAGAACTCTCCGACGCTGTCAAACGAGGCGGTAGCC---CAGCTTAT 523
Db 139 pAlaSerAspValGluValLeuGlyProLeuArgProGlyValAlaGlySerAlaLe 159
Qy 522 CATTCAATAACTCCGGCAGAACTTTGAAAGGTCCCTCGGTAGCTCATGTTCCGACAG 463
Db 159 uValGlnValArgValArgGluLeuArgGlyGlyGluAlaGluArgGlyGlyAlaGly 179
Qy 462 AGCGGTGGAAGTATAATA-----AGACACACCACG-- 432
Db 179 yGlyGlyLeuPheSerLeuCysAlaTrpAspGlyArgAlaTrpHisHisGlyAl 199
Qy 431 -----ATCCTTGGGTTTGGCGACCATATAGATACCTTGGAGTCTCCATTGT 385
Db 199 aAlaGlyGlyPheLeuLeuArgValArgProArgLeuTrpGlyProGlyGlyAspLeuLe 219
Qy 384 GGAGGTCAGGGCAGAAAGCGGTAGGCGAG-----GTGGCTCTTGATAGGCGGG 334
Db 219 uProProAlaTrpLeuArgAlaLeuGlyAlaLeuLeuLeuLeuAlaLeuSerAlaLeuPh 239
Qy 333 GAGAGGT-----CAGTCTTGATGATGTCAGGCGGCTCATCGGGGTGAGCGGCT 289
Db 239 eSerGlyLeuArgLeuSerLeuLeuSerLeuAspProValGluLeuArg-----ValLe 257
Qy 288 CCAGACCGGAGTGTGCTCGTGTGATGTCATCAAGCCGATCTCATCGGGGTGAGCGGCT 229
Db 257 uArgAsnSerGlySerAlaAlaGluGlnGluGlnAlaArg-----ArgValGlnAlaVa 275
Qy 228 GCCTCACCAGTAGACAC-----CTCCTGACGAGAGCTGTGGTGGTACT 184
Db 275 lArgGlyArgGlyThrHisLeuLeuLeuLeuLeuLeuLeuGlyGlnAlaGlyAlaAsnAl 295
```

```
183 T-----GGGGTAGGTGAC-----GA 169
295 aAlaLeuAlaGlyTrpLeuTrpAlaSerLeuProGlyValGlyAspProGlyGluAs 315
Qy 168 TCACACGTCGCTGGCGGCAC 147
Db 315 pSerGlyGluAlaGlyValHis 322

RESULT 30
US-10-505-928-150
; Sequence 150, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 150
; LENGTH: 5738
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-150

Alignment Scores:
Pred. No.: 1.72 Length: 5738
Score: 95.00 Matches: 152
Percent Similarity: 29.3% Conservative: 76
Best Local Similarity: 19.6% Mismatches: 293
Query Match: 2.2% Indels: 257
DB: 6 Gaps: 31

US-10-768-158-1 (1-2419) x US-10-505-928-150 (1-5738)
Qy 94 TCGGCTGCGCCCTTCTGCGCGGGAAGATGGAGGAGATCGCCAACTTCC----- 144
Db 4211 CysAspCysThrAspAlaGlnGlyHisSerTrp-----AlaProGlySerGlnHisGln 4228
Qy 145 -----CGGTGGCGCCAGCGACGTGTGGATCGTCACTACC 180
Db 4229 AspAlaCysAsnAenCysSerCysGlnAlaGlyGlnLeuSerCysThrAlaGlnProCys 4248
Qy 181 CCAAGTCGGCACCAGCTTGTGACGAGGTGTCTACTTGTGTGAGCCAGGCGGCTGACC 240
Db 4249 ProProProThrHisCysAlaTrpSerHisTrpSerAlaTrp-----Ser 4263
Qy 241 CCGATGAGATCGGCTTGTGATGAACATCGACGAGCAGCTCCCGGTCTCGAGTACCACAGC 300
Db 4264 ProCys-----SerHisSerCysGlyProArgGlyGlnGlnSerArgPhe 4278
Qy 301 CGGCGCC-----TGGACATCATCAAGAACTGACCTCTCCCC-----GCCTCATCA 345
Db 4279 ArgCysGlyProGlyLeuAlaSerArgSerCysProCysLeuMetAlaLysAla 4298
Qy 346 AGAGCCACCTGCCCTACCGCTTTC----- 369
Db 4299 AspProThrCysAsnSerThrPheLeuHisLeuAspThrGlnGlyCysTyrSerGlyPro 4318
Qy 370 TGCCTCTGACCTCCCAATGAGACATCCCAAGGTCACTATATGCTCGCAACCCCAAGG 429
Db 4319 CysProAspSerCys-GlnTrpSerLeuTrpGlyProTrpSerProCysGlnValProCys 4338
Qy 430 ATCTGGTGGTCTCTATTATC-----AGTTCACCGCTCTCTCGGACCATGAGCTTAC 482
Db 4338 sSerGlyGlyPheArgLeuArgTrpArgGluAlaGluAlaLeuCysGlyGlyGlyCysAr 4358
Qy 483 CGAGGCACCTTTCAAGAAATTCGCGGAGGTTTATGAATGATAGCTGGGCTCGGCTCC 542
Db 4358 gGlu-----ProTrpAlaGlnAspAr 4365
```


QY 543 TGGTTTG-----AGCAGTGCAGGAGTTCTCGGAGCACCGCATGGAC 584
Db 4365 glysLeuGlnArgArgAlaLeuProSerThrCysValAsnGluSerLeuValCysProHi 4385
QY 585 TCGAACGTGCTTTTCTCAAGTATGAACACATGCATCGGACCTGGTGCACGATGGTGAG 644
Db 4385 sGlnGluCys-----ProValLeuGlyProTrpSerAlaTrpSerSe 4399
QY 645 CAGCTGGCCAGATTCTCGGGGTGCTCTGCACAGG-----CC 683
Db 4399 rCysSerAlaProCysGlyGlyThrMetGluArgHisargThrCysGluGlyGlyPr 4419
QY 684 CAGCTGGAAGCCCTGACGAGACACTGCCACAGCTGGTGACCATGCTGCAACCTGAG 743
Db 4419 oGlyValAlaProCysGlnAlaGlnAspThrGlu---GlnArgGlnGluCysAsnLeuG1 4438
QY 744 GCCCTGCCCGTGGGCC-----GGGGAAGATTGGCTGTGGAGACATCTTACCG-- 795
Db 4438 nProCysProGluCysProProGlyGlnValLeuSerAlaCysAlaThrSerCysProCy 4458
QY 795 ----- 795
Db 4458 sLeuCysTrpHisLeuGlnProGlyAlaIleCysValGlnGluProCysGlnProGlyCy 4478
QY 796 -----TCTCCATGAATGAGAAGTTTGACTTGTGTATTAACAGAAAGATGGGAAGTGT 848
Db 4478 sGlyCysProGlyGlnHisSerLeuProTrpGlyLeuThrLeuGluGluG1 4498
QY 849 GACCTCAGTTTGACT-----TTTATTATTAATAACAGAAACACACACCTGCATCTCA 902
Db 4498 nAlaGlnGluLeuProProGlyThrVal-LeuThrArgAsnCysThrArgCysValCysH 4518
QY 903 CAATACCCAGACAGTCTACTAGCCAAAGTCTGTATGCATTTATTTCTTCTGCTGGA 962
Db 4518 isGlyGlyAlaPheSerCysSerLeuValAspCysGlnGlyLeuValPro-ProGly 4537
QY 963 CAAACTCTGGAAGCA----- 977
Db 4538 GluThrTrpGlnGlnValAlaProGlyGluLeuGlyLeuCysGluGlnThrCysLeuGlu 4557
QY 978 -----GGGTGTAACAGCGGGGAAGGAGCGCGTGCAGCGGAGGAGTGTGATG 1031
Db 4558 MetAsnAlaThrLysThrGlnSerAsnCysSerSerAlaArgAlaSerGlyCysValCys 4577
QY 1032 ATTCCTCAACCGAAGCAGCTGTCTCGCTTTAGAACGTGCAGCTCTCCATCTCTGATTAC 1091
Db 4578 GlnProGlyHis-----PheArgSer-GlnAlaGlyProCysValProG1 4592
QY 1092 AAACAGTCTCCACATTGCAGTTCCAAATGGCTGGACCGTAAGGATAAAGCTGTAATATA 1151
Db 4592 uAspHisCysGlu-----CysTrpHisLeuGlyArgProHisLeuValAr 4607
QY 1152 TGCACCTAGATGTCTGCTTTTCAACCCCGTATTATTGATTTATAGACCTTTTCACT 1211
Db 4607 gHisArgThrProSerAlaThrHisProPheLeu-----ThrProSerLe 4623
QY 1212 GGAATCTACATAAATGTCAAGTAAACCAAAATAAAGTTTCATTTCCAGGGGAATCAGGAG 1271
Db 4623 uProIleCysLeuGlyArgGlyProGluProTrpLeuGlySerValGlnGlyGlnArgSe 4643
QY 1272 C-----GAGCCACACCCGAATGTGTAGAAAGACTCTCAGGGTTAACTCTTTATTTTGTAG 1325
Db 4643 rLeuLeuGluHisHisPro---GlyLysLys----- 4652
QY 1326 TTTTATTATCTAAGGCACAGCCATCTGTTCTCACTGGTTCTGAGATAGTGGTGAGAAC 1385
Db 4652 ----- 4652
QY 1386 AGAGGATGAGTTGGTCTGTTGGGGGAATCTGGACACTGTTGTTATTCTGACGGAGTTCA 1445
Db 4653 -----GlnMetGlyTyrMetGly----- 4658

QY 1446 CTTCTTCAGAACCTTCTCTGAAATGAGCAGAAATGTTTCACTAGGTCTTTCAGAAATGGACGT 1505
Db 4659 -----GluMetGluVa 4662
QY 1506 CTTTCTGCAGAGACTTCCAGCGGCGCTCCAAAGGCCCAATGCAGAGGAGCGCGGA 1565
Db 4662 lGlnGlyProThrArgGluSerGlyGlnSerLeuProProGln----- 4676
QY 1566 GCATGTGCTGAGGAAGTCTGCTGCTGGTGGAGCTGCGAGGTGGAGTCTTAATGTCAGTCAGG 1625
Db 4677 -----LysLysAlaTyrLeuSerHisLeuSerThrGlySerGlyHisIleGluG1 4693
QY 1626 AGCATTTGCATGCGAGTGGTGGAGAGTGGCGCCACCAAGAGCCAGTTCCTCGGAATT 1685
Db 4693 yAspTrpAlaGlyArgAsnArgLysLeuLeuLysProArgSerIleGlnLysSerTrpPh 4713
QY 1686 T-----GAGCTGAATTCACACAGCCTTACTTTGTTTCTCTGAA 1721
Db 4713 eValGlnPheProTrpLeulleMetAsnGluGluGlnThrAlaLeuPheCys----- 4730
QY 1722 GTGATAGCTACTAATGCTGCAAGCAGATGCTTTAATAGTAAATTTCTAAATATCCCGGG 1781
Db 4731 -----SerAlaCysArgGluTyrProSerIleArgAspLysArgSerArgLe 4746
QY 1782 TCTT-----TATCATTCAGTTTGTTC 1802
Db 4746 uIleGluGlyTyrThrGlyProPheLysValGluThrLeuLysTyrHisAlaLysSerLy 4766
QY 1803 TGTCCACTGAGGCGCTCAGCGTGGGAGGACCAATTTTGGAGTGTAGCCCTGTTTCACT 1862
Db 4766 sAlaHisMet-----PheCysValAsnAlaLeuAlaAlaAr 4778
QY 1863 CGATCAGTTTGGCAGCGCGCTGCTGTCTGTCCACCTCATCCCTCCGTGTATCTGAG 1922
Db 4778 gAspProIleTrpAlaalaargPheArg-SerIleArgAspProProGlyAspValLeuA 4798
QY 1923 GGAGTAAAGTGAGTCTTTATGCTTCACCTGCTCACTGCTCAATTTCTCACCC----- 1970
Db 4798 laSerProGluProPheThrAlaAsp-CysProIlePheTyrProProGlyProLeu 4817
QY 1971 -----ACATTCTGCTGAAGCGATGAGAGTGGGGCCGAGTAGCCAGCCACACCC 2018
Db 4818 GlyGlyPheAspSerMetAlaGluLeuLeuProSerSerArgAlaGluLeuGluAspPro 4837
QY 2019 CGTGGGACCGGGTGTCTCTCATTTATGTTGGTGGAA 2057
Db 4838 GlyGlyAsp---GlyAlaIleProAlaMetTyrLeuAsp 4849

RESULT 31

US-10-196-749-284
; Sequence 284, Application US/10196749
; Publication No. US20060094864A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C340
; CURRENT APPLICATION NUMBER: US/10/196,749
; PRIOR FILING DATE: 2002-07-16
; PRIOR FILING DATE: 2002-01-15
; PRIOR FILING DATE: 1997-09-18


```

Db      382 rserArgGlyProSerAspThrSerSerArgPheSerLysGluGlnArgGlyValAlaG1 402
Qy      674 -----TGCAAGGCCCCAGCTGGGAAGCCCTGACGG 702
Db      402 nProLeuAlaGlnLeuArgSerCysProGlnGluGlyProArgGlyArgGlyLeuAl 422
Qy      703 AGCACTGCCA-----CCAGCTGGTGGACCACTGCTGCACCGCTGAG----- 743
Db      422 aAlaArgProLeuGluGlnArgAlaGlyGlyProVal-AlaArgSerGluGluProGlyA 442
Qy      744 --GCCTCCCTGGCGCCGGGAGAGTT-----GGGCTGTGGGAAGGACATCTTCA 792
Db      442 laProLeuProValAlaValGlyThrAlaGluProGlyGlySerMetLysThrThrPheT 462
Qy      793 CCGTCTCCATGAATGAG 809
Db      462 hrIleGluIleLysAsp 467

RESULT 33
US-10-525-621-1
; Sequence 1, Application US/10525621
; Publication No. US20060100418A1
; GENERAL INFORMATION:
; APPLICANT: Kiyosue, Yuko
; APPLICANT: Sasaki, Hiroyuki
; APPLICANT: Tsukita, Shoichiro
; APPLICANT: Eisai Co., Ltd.
; TITLE OF INVENTION: CULTURED XENOPUS LAEVIS CELL LINES
; TITLE OF INVENTION: EXPRESSING MUTANT ADENOMATOUS POLYPOSIS COLI GENE
; FILE REFERENCE: 082368-002400US
; CURRENT APPLICATION NUMBER: US/10/525,621
; PRIOR FILING DATE: 2005-02-22
; PRIOR APPLICATION NUMBER: PCT/JP03/10434
; PRIOR FILING DATE: 2003-08-19
; PRIOR APPLICATION NUMBER: JP 2002-241487
; PRIOR FILING DATE: 2002-08-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2829
; TYPE: PRT
; ORGANISM: Xenopus laevis
US-10-525-621-1

Alignment Scores:
Pred. No.: 1.73 Length: 2829
Score: 93.50 Matches: 82
Percent Similarity: 32.0% Conservative: 26
Best Local Similarity: 24.3% Mismatches: 140
Query Match: 2.1% Indels: 89
DB: 6 Gaps: 15

US-10-768-158-1 (1-2419) x US-10-525-621-1 (1-2829)
Qy      840 CCAATCTTCTGTTATACACCAAGTCAACTTCTCATTCATGAGACGGTGAAGATGCTCT 781
Db      2286 ProSerSer-----ThrProGlyCysSerLysGlyAsnSerArgSerGlySerArg 2303
Qy      780 TCCACAGCCCAACTCTTCCCGGCCCCACGG-----GCAGGGCCCTCAG 739
Db      2304 AspSerAlaSerArgProSerProGlnProLeuSerArgProLeuGlnSerProGly 2323
Qy      738 CGTTGCAGCACTGGTCCACCAAGCTGGTGCAGTGTCCG-----TCAGGGGCTT 691
Db      2324 ArgAsnSerIleSerProGlyLysAsnGlyLysSerProAsnLysPheSerGlnLeu 2343
Qy      690 CCA-----GCTGGGCTTGTGCAC 673
Db      2344 ProArgThrThrSerProSerThrAlaSerThrLysSerSerGlySerGlyArgMetSer 2363
Qy      672 AGGACACCCCGAGGAACTGGCCAGCTGCTCA----- 640

```

```

Db      2364 TyrThrSerProGlyArgGlnLeuSerGlnProAsnLeuSerLysGlnSerGlyLeuPro 2383
Qy      639 -----CCATCTCTACCAAGTCCCGATGCATGCTCTTCAT----- 607
Db      2384 LysThrHisSerSerIleProArgSerGluSerAlaSerLysSerLeuAsnGlnAsnVal 2403
Qy      606 -----ACTTCAGAAAAGACACGTCGAGTCCATGCGGTGCTCCAGCAACTCCCTGCA 556
Db      2404 AsnThrGlySerAsnLysLysValGluLeuSerArgMetSerSerThrLysSerSerGly 2423
Qy      555 CGTCTCAAAACGAGGACCGTAGCCACGCT-----TATCATTCATAA 514
Db      2424 SerGluSerAspArgSerGluArgProAlaLeuValArgGlnSerThrPheIleLysGlu 2443
Qy      513 ACCTCCGGCAGAAATTCGTGAAGTGCCTCGGTAGTCATGGTCCGAGAGAGCGGTGGA 454
Db      2444 AlaProSerProThrLeuArgArgLysLeuGluSerAlaSerPheGluSerLeuSer 2463
Qy      453 ACTGATAATAAG-----ACACCAACAGATCCTTGGGGTTCGAGGCCATATAGATGACCT 400
Db      2464 SerSerSerArgAlaAspSerProProArgSerGlnThrGlnThrProAlaLeuSerPro 2483
Qy      399 TGGAGTCTCCATTCGTGGAGGTACAGGGGCAGAAAGCGGTAGGCGAGGTGGCTCTTGATGA 340
Db      2484 ---SerLeuProAspMetAlaLeuSerThrHisSerIleGlnAla----- 2497
Qy      339 GCGCGGAGAGGTGAGTTCCTTGATGATGTCAGGCCGCGGTGGGTACTCCAGGACCG 280
Db      2498 GlyGlyTrpArgLysMetPro-----ProAsnLeuAsnProAlaAlaGluHisGly 2514
Qy      279 GGACCTCTCTCGATGTTTCATCAAGCGGATCTCATCGGGTTCAGCCCTGGTGGTCCACCA 220
Db      2515 AspSerArgArgArgHisAspIleSerArgSerHis-----SerGluSerPro 2530
Qy      219 AGTAGACCACTCTCTCGACCAAGCTGTCGCGGACTTGGGTAGGTGACGATCCACACGT 160
Db      2531 SerArgLeuProlleThrArgSer-----GlyThrTrpLysArgGluHisSerLysHis 2548
Qy      159 CGCTGGGCCGACCG-----GGAAGTGGCGATCTCTCCATCTCCCGCGGAGAGG 106
Db      2549 SerSerSerLeuProArgValSerThrTrpArgArgThrGlySerSerSerIleLeu 2568
Qy      105 GCGGCAGCGCACGCCAT----- 88
Db      2569 SerAlaSerSerGluSerSerGluLysAlaLysSerGluAspGluLysGlnGlnValCys 2589
Qy      87 -----GGAAGTTCGAAGTACTTCTCGAACTCCCGCGGTGTCTGG 46
Db      2589 SerPheProGlyProArgSerGluCysSerSerSerAlaLysGlyThrTrp 2605

RESULT 34
US-10-505-928-851
; Sequence 851, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 851
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-851

Alignment Scores:
Pred. No.: 0.847 Length: 353
Score: 93.00 Matches: 81

```

```
Percent Similarity: 29.7%      Conservative: 24
Best Local Similarity: 22.8%    Mismatches: 113
Query Match: 6                 Indels: 135
DB:                             Gaps: 18

US-10-768-158-1 (1-2419) x US-10-505-928-851 (1-1353)

QY 1 CGGACGGCGGCGGGCGGCGATGG-----CGGAGAGCGAGCGCG 39
D 15 AlaArgProAnaAlaGlyAlaTrpSerIleGluAlaGlyProSerArgAspSerArgAla 34
QY 40 AGACCCCGCAGCA-----CCCGCGGGAGTTCGAGAGCAAGTACCTCGAGT 84
D 35 TrpProAlaAlaGlyAsnHisThrLeuAlaGlnThrAlaSerProAspGlyThrGluPro 54
QY 85 TCCATGGCGTGGCGTGGCGGCTTCTGCCCGGGAAGATGAGGAGATGCCAACT--- 141
D 55 GlyHisSerProGlyCysLeuAsnSerAlaGly-----ValSerThrProThrSer 71
QY 142 ---TCCCGGTGGCGGCCA----- 156
D 72 GlySerMetCysProProGlnAlaGlnAlaGluValGlyProThrMetThrGluLeuAla 91
QY 157 ---GCGAGCTGTGGATCGTCACTACCCAGTCCGCGCACCAGCTTGCTGCAGGAGGTG 213
D 92 GluMetValCysAlaProSerProAlaProAlaProProProlys-ProAla----- 108
QY 214 TCTACTTGGTGGAGCGCGCTGACCCGATGATCGGCTTGATGAACATCGACGAGC 273
D 109 -----SerProGlyPro---ProGlnValGluGluValGlyHisArgGlyGlySe 124
QY 274 AGCTTCCC-----GGTCTGGAGTACCCACGCGGGCTGGACATCATCAAGNAC 324
D 124 rSerProProArgLeuProProGlyValProValIleSerLeuGlyHisSerArgProPr 144
QY 325 TGACCTTCCCGCTCATCAAGAG-----CCACCTGCCCTACCGCTTTC 369
D 144 oGlyValAlaMetProThrThrGluLeuGlyThrLeuArgProProLeuLeuGlnLeuSe 164
QY 370 TGCCCTC-----TGACCTCCCAATGAGACTCCAAGTCACTATAT----- 412
D 164 rThrLeuGlyThrAlaProProThrLeuAlaLeuHisThrHisProHisProPheLeuAs 184
QY 412 ----- 412
D 184 nSerValTyrIleGlyProAlaGlyProPheSerIlePheProSerSerArgLeuLysAr 204
QY 413 -----GGCTCGCAACCC----- 424
D 204 gArgProSerHisCysGluLeuAspLeuAlaGluGlyHisGlnProGlnLysValAlaAr 224
QY 425 -----CAAGGATCTGGTGGTGTCTTATT 447
D 224 gArgValPheThrAsnSerArgGluArgTrpArgGlnGlnAsnValAsnGlyAlaPheAl 244
QY 448 ATCAGTTCACCGCTCTCTGGGACCATGAGTACCGAGGACCTTTCAAGAACTCTGCC 507
D 244 aGluLeuArgLysLeuLeuProThrHisProProAspArgLysLeuSerLys-----As 262
QY 508 GGAGTTTATGAATGAATGAAGCTGGGCTACGGCTC----- 541
D 262 nGluVal-----LeuArgLeuAlaMetLysTyrIleGlyPheLeuVa 276
QY 542 ---CTGTTTGTAGCACTGTGAGGAGTCTGGGAGCA-----CCGATGGA----- 583
D 276 lArgLeuLeuArgAspGlnAlaAlaAlaAlaAlaGlyProThrProGlyProAr 296
QY 584 -----CTCGAACGTGCTTTTCTCAAGTATGAAGACATGATCGGACCTGGTGA 633
D 296 gLysArgProValHisAsa:gvalProAspAspGlyProArgArgGlySerGlyArgArgAl 316
QY 634 CGATGGTGGAGCAGCTGGCGCAGATTCTTGGGGGTGTCTGTGTGACAAGGCCCACTGGAAG 693
D 316 aGluAlaAlaAlaArgSerGln-----ProAlaProProAlaAs 329
QY 694 CCCTGACGGAGCACTGCCACCACTGGTGGACCACTGGTGGACCACTGGTGG 730
D 329 pProAspGlySerProGlyGlyAlaAlaArgProIle 341

RESULT 35
US-10-511-937-2466
; Sequence 2466, Application US/10511937
; Publication No. US2006008836A1
; GENERAL INFORMATION:
; APPLICANT: EXPRESSION DIAGNOSTICS, INC.
; APPLICANT: Wohlgemuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; APPLICANT: Prentice, James
; APPLICANT: Morris, MacDonald
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND MONITORING TRANSPLANT REJECTION
; FILE REFERENCE: 506612000104
; CURRENT APPLICATION NUMBER: US/10/511.937
; PRIOR FILING DATE: 2004-10-19
; PRIOR APPLICATION NUMBER: PCT/US2003/012946
; PRIOR FILING DATE: 2003-04-24
; PRIOR APPLICATION NUMBER: US 10/131,831
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US 10/325,899
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 3117
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2466
; LENGTH: 362
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-511-937-2466

Alignment Scores:
Pred. No.: 0.855      Length: 362
Score: 93.00         Matches: 67
Percent Similarity: 30.4%      Conservative: 20
Best Local Similarity: 23.4%    Mismatches: 133
Query Match: 2.1%           Indels: 66
DB:                          Gaps: 12

US-10-768-158-1 (1-2419) x US-10-511-937-2466 (1-362)

QY 765 TTCCCGCGCCCGCAGCGGAGG-----CCTCAGGTTTGCAGCACTGGTCCACCA----- 718
D 82 PheSerGlyProGluProGlyGlyAlaProGlnThrCysAlaLeuAlaProSerGluAla 101
QY 717 -----GCTGGTGGCAGTGCT 703
D 102 SerGlyAlaGlnTyrProProProGluThrLeuGlyAlaTyrAlaGlyGly----- 119
QY 702 CCCTCAGGCTTCCAGCTGGGCTTGTACAGGACACCCCGAGGAACTGCCAGCTGCT 643
D 120 ---ProGlyLeuValAlaGlyLeuLeuGlySerGluAspHisSerGlyTrpValArgPro 138
QY 642 CCACCATCGTCACGAGTCCGATGCTTCTATCTTGCAGAAAGACGCTTCGAGT 583
D 139 AlaLeuArgAlaArgAlaProAspAlaPheVal-----Gly 150
QY 582 CCATGCGGTGCTCCAGAACTCTCGACAGTCTCAACAGGAGCCGTAGCCAGCTTAT 523
D 151 ProAlaLeuAlaProAlaProAlaProGluProLysAlaLeuAlaLeuGlnProValTyr 170
QY 522 -----CATTATAACCTCCGGAGATTCTTGAAGGTGCTCGGTAGCTCATGG 472
D 171 ProGlyProGlyAlaGlySerSerGlyGlyTyrPheProArg---ThrGlyLeuSerVal 189
QY 471 TCCGAGAGAGCGGTGGAACGTATAATAAGACACCAACAGATCCTTGGGGTTGC----- 418
```

```

Db      190 ProAlaAlaSerGlyAla-----ProTyrGlyLeuLeuSer 201
Qy      417 ---GAGCCATATAGATGACCTTGGAGTCTCCATTCTGGAGGTTCAGAGGCGCAGAAAGCGGT 361
Db      202 GlyTyrProAlaMetTyrProAlaProGlnTyrGlnGlyHisPheGlnLeuPheArgGly 221
Qy      360 AGGGCAGGTGGCTCTTGATGAGCGGGGAGAGGTCACTTCTTGATGATGTCAGGCCCG 301
Db      222 LeuGlnGlyProAla-----ProGlyPro 229
Qy      300 GCTGTGGGTACTCCA-----GGACCGGAGCTGCTCGTCGATGTTCA 256
Db      230 AlaThrSerProSerPheLeuSerCysLeuGlyProGlyThrVal-----GlyThr 246
Qy      255 AGCCGATCTCATCGGGGTACAGCGCTCGCTCCACCAAGTAGACCACTCTCTGCAGCAAGC 196
Db      247 GlyLeuGlyGlyThrAlaGluAspProGlyValIleAlaGluThrAlaProSerIysArg 266
Qy      195 TGGTCCCGACTTGGGTAGTGACGATCCACACGCTCGTGGCGGCA---CCGGGAAGT 139
Db      267 GlyArgArgSerTrpAlaArgIysArgGlnAlaAlaHisThrCysAlaHisProGlyCys 286
Qy      138 TGGCGATCTCTCCATCTTCCCGCGGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 79
Db      287 GlyIysSerTyrThrLysSerSerHisLeuIysAlaHisLeuArgThrHisThrGlyGlu 306
Qy      78 AGTACTCTCTCGAAGTCCCGCGGGGTGCTGGGGGTCTCGGCTCTCGCTCTCGGCGGCGG 19
Db      307 LysProTyrAlaCysThrTrpGluGlyCysGlyTrpArgPheAlaArgSerAspGluLeu 326
Qy      18 CGCGCGCTCGCGCTCGC 1
Db      327 ThrArgHisTyrArgLys 332

RESULT 36
US-10-196-749-238
; Sequence 238, Application US/10196749
; Publication No. US20060094864A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C340
; CURRENT APPLICATION NUMBER: US/10/196,749
; CURRENT FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541

```

```

; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 238
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-196-749-238

Alignment Scores:
Pred. No.: 0.903      Length: 325
Score: 92.50         Matches: 62
Percent Similarity: 32.0%      Conservative: 24
Best Local Similarity: 23.0%    Mismatches: 72
Query Match: 2.1%             Indels: 111
DB: 6                    Gaps: 16

US-10-768-158-1 (1-2419) x US-10-196-749-238 (1-325)
Qy      86 CCATGGGCTCGGCTCGGCTCGCCCTTCTCGCGGGAAGATGGAGGATCGGCAACTTCCC 145
Db      115 ProTrpGlnMetMetAlaAla-----121
Qy      146 GGTCCGCGCCAGCAGCGTGTGGATGTCACCTACCCCAAGTCCGGCACCAGCTTGTGCA 205
Db      122 ---AlaAlaGluAspArgLeuGlyGluAlaLeuProGluLeuSerTyrLeuSerSer 140
Qy      206 GGAGGTGGTCTACTTGGTGAGCCAGGGCGGTGACCCCGATGAGATCGGCTTGTGAACAT 265
Db      141 AlaAlaAlaLeu-----144
Qy      266 CGACGAGCAGCTCCCGGTCTGTGGAGTACCCACAGCGGCGCTGGACATCATCAAGAACT 325
Db      145 -----AlaProGlySerGly---ProLeuProGlyGluSerSerProAspAlaThr 160
Qy      326 GACCTCTCCCGCT-----CATCAAGAGCCACCTGCCTACCGCTTCTTCTGCC 373
Db      161 GlyLeuSerProGluAlaSerLeuLeuHisGlnAspSerGluSerArgArgLeu----- 178
Qy      374 CTCTGACCTCCACAATGGAGACTCCAAGGTCTATATATGCTGCGAACCTGCTTTTCAAGTA 433
Db      179 -----ProArg-----SerAsnSerLeuGlyAla 186
Qy      434 GGTGGTGTCTATTATCAGTT---CCACCGCTCTCTCGGACCATGAG---CTACCCGAGG 487
Db      187 GlyGlyIysIleLeuSerGlnArgProTrpSerLeuIleHisArgValLeuProAsp 206
Qy      488 CACCTT-----493
Db      207 HisProTrpGlyThrLeuAsnProSerValSerTrpGlyGlyGlyProGlyThrGly 226
Qy      494 -----TCAAGAATTCTCGCGAGGTTTATGAATGATAAGCTGGCTACGGCTCTCGTT 547
Db      227 TrpGlyThrArgProMetProHisPro-----GluGlyIle-----238
Qy      548 TGACGACGTGCAGGATTTCTGGGAGCACCGCATGGACTCGAACCTGCTTTTCTCAAGTA 607
Db      239 -----TrpGlyIleAsnAsnGlnProProGly-----ThrSerTrpGlyAsnIle 253
Qy      608 TGAAGACATGCATCGGAGCTGTGTGACGATGGTGCAGCAGCTGCCAGATTCCTGGGGGT 667
Db      254 AsnArgTyr-----ProGlyGlySerTrpGlyAsnIleAsnArgTyrProGlyGly 270
Qy      668 GTC-----CTGTGACAAGGCCAGCTGGAGC-----694
Db      271 SerTrpGlyAsnIleAsnArgTyrProGlyGlySerTrpGlyAsnIleHisLeuTyrPro 290
Qy      695 ---CCTGACGAGCACCTGCCACACAGCTGGTGGACAGTGTGCAACGCTGAGGCGCTGCC 751
Db      291 GlyIleAsnAsnProPheProPro-----GlyValLeu 301

```



```

; APPLICANT: Lee, Nancy M
; TITLE OF INVENTION: DRUG SCREENING AND MOLECULAR DIAGNOSTIC TEST FOR EARLY DETECTION
; FILE OF INVENTION: OF COLORECTAL CANCER: REAGENTS, METHODS, AND KITS THEREOF
; FILE REFERENCE: NLEE-01001US1 MCF/MLB
; CURRENT APPLICATION NUMBER: US/11/242,111
; CURRENT FILING DATE: 2005-09-29
; PRIOR APPLICATION NUMBER: 60/614,746
; PRIOR FILING DATE: 2004-09-30
; PRIOR APPLICATION NUMBER: 60/651,344
; PRIOR FILING DATE: 2005-02-08
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 20
; LENGTH: 554
; TYPE: PRT
; ORGANISM: HUMAN
US-11-242-111-20

Alignment Scores:
Pred. No.: 1.35 Length: 554
Score: 91.50 Matches: 74
Percent Similarity: 30.5% Conservative: 22
Best Local Similarity: 23.5% Mismatches: 129
Query Match: 2.1% Indels: 90
DB: 7 Gaps: 18

US-10-768-158-1 (1-2419) x US-11-242-111-20 (1-554)
QY 832 TGTATACACCAAGTCAAACTCTCATTCATGAGACGGTGAAGATGCTCTCCACAGC 773
DB 191 CysLeuYrProLysAla---IleProSerSerAspProAlaSerValSerProHisGln 209
QY 772 CCAACTCTTCCCGGCCCCACGGCAGGCGCTCAGCGTTGCAGCACTGTCACACAGTGG 713
DB 210 ProLeuAlaProSerMetAlaProValAlaGlyLeuThr-----TipGluAspSer-Gl 227
QY 712 TGGCAGTCTCGGCAGCGGCTCCAGCTGG-----CCTGTGCACAGCACACCCAGG 659
DB 227 uGlyThrGluGlySerSerLeuLeuProGlyGluGlnProLeuHisThrValAspProGl 247
QY 658 AATCTGGCCAGCTGCTCCACCATCGTCA-----CCAGGTCGCCGATGCATG 614
DB 247 YSerAlaLysGlnArgProProArgSerThrCysGlnSerPheGluProProGluThrPr 267
QY 613 TCTTCATACTTGAGAAAAGCAGCGTTCGAGTCCATCGCGTGTCCAGAACTCTGACAG 554
DB 267 oVal-----ValLysAspSerThrIleGlyGlySerProGlnPro--Ar 281
QY 553 TGCTCAACACAGGACCGTAGCCAGCTTATCATTTCAAACTCCGCGAGAATTTCTGA 494
DB 281 gProSerValGlyAlaPheAsnProGlyMetGluAsp-----IleLeuAs 296
QY 493 AAGGTGCTCGGTAGCTCATGTGTCGCGAGAGCGGTGGAACCTGATAATAAGACACACC 434
DB 296 pSerAlaMetGlyThrAsnTrpValProGluGluAlaSerGlyGluAlaSerGluIlePr 316
QY 433 AGATCCTTGGGGTTCGAG-----CCATATAGATGACCTTGGAGTCTCCATTTGGGAGG 380
DB 316 oValProGlnGlyThrGluLeuSerProSerArg--ProGlyGly-----Glyse 332
QY 379 TCAGAGGCGCAGAACCGGTAGGCGAGGTGGCTCTTGATGAGCGGGGAGAGGTTCAGTCC 320
DB 332 rMetGlnThrGluProAlaArgProSerAsnPhe-----LeuSerAlaSerSerPr 349
QY 319 TTGATGATGTCAGGCGCGGTGTGGGT-----AC 290
DB 349 o-----LeuProAlaSerAlaLysGlyGlnGlnProAlaAspValThrAlaThrAlaLe 367
QY 289 TCCAGGACCGGAGGTGCTGTCGATGTTTCATCAAGCGCATCTCATCGGGGTACGCGCC 230
DB 367 uProArgValGlyProValMetProThrGlyGlnAspTrpAsnHisThrProGlnLysTh 387
QY 229 TGGCTCACCAAGT-----AGACCACCTCTCTCAGCAAGCTGGTCCCGAGCTTGGG 179
```

```

DB 387 rAspHisProSerAlaLeuLeuArgAspProPro----- 398
QY 178 TAGGTGACGATCCACACGTCGCTGGCGCCGACCGGGAAGTTGGCGA----- 133
DB 399 -----GluProGlySerProArgIleSerSerLeuAr 409
QY 132 -----TCTCTCATCTTCCCGGGCAGAGGGCGGCACCGCACGC----- 91
DB 409 gProGlnAlaLeuSerAsnProSerThrLeuSerAlaGlnProGlnLeuSerArgSerHi 429
QY 90 -----CATGGAACCTCAAAGTACTTCTCTCTCG 65
DB 429 sSerSerGlySerValLeuProLeuGlyGluLeuGluGlyArgArgSerThrArgAspAr 449
QY 64 AACTCCCCCGGGTGTGGGGGTCTCGGCTCTCGCTCTCCGCCA 22
DB 449 gThrSerPro-----AlaGluProGluAlaAlaPro 459

RESULT 41
US-10-322-836-48
; Sequence 48, Application US/10322836
; Publication No. US2006090212A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Alexander C., Jr.
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Freidrich, Glenn A.
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: NOVEL HUMAN GENES AND PROTEINS
; TITLE OF INVENTION: ENCODED THEREBY
; FILE REFERENCE: 8535-0037-999
; CURRENT APPLICATION NUMBER: US/10/322,836
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: US/09/579,114
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 48
; LENGTH: 1075
; TYPE: PRT
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (130)...(130)
; OTHER INFORMATION: Xaa = Gln or STOP
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (179)...(179)
; OTHER INFORMATION: Xaa = Ser or Gly
US-10-322-836-48

Alignment Scores:
Pred. No.: 1.74 Length: 1075
Score: 91.50 Matches: 74
Percent Similarity: 31.9% Conservative: 23
Best Local Similarity: 24.3% Mismatches: 89
Query Match: 2.1% Indels: 118
DB: 6 Gaps: 16

US-10-768-158-1 (1-2419) x US-10-322-836-48 (1-1075)
QY 768 CTCCTCCCGGCCCCACCG-----GCAGGGGCT-----CAGCGTTCGACACTGGTCCA 721
DB 669 IlePheProSerProArgGluLeuGluGlyProValTyGluLysCysMet----- 685
QY 720 CCAGCTGTGGCAGTGTCCGTCAGGGCTTCAGCTGGCGCTTGTCCAGGACACCCCCA 661
DB 686 ---AlaGlyGlyGluGluTyrcysAspSerPro-----HisSerGluProGly 700
QY 660 GGAATCTGGCCAGCTGTCTCCACCATCGCTCACAGGTCCCGATGCGTCTTCATCTTGA 601
DB 701 AlaIle-AspGluValAspHisAsnGlyThrGluPro---HisThrSerAspGluGl 719
```



```
DB: 6 Gaps: 10
US-10-768-158-1 (1-2419) x US-10-509-131-25 (1-783)
QY 786 TGTCTCTCCACAGCCAACTCTTCCCGCCGCGGCGGCGG-----CCT 742
Db 633 CysProSerThrSerSerLeuPheAsnLeuGlnLysSerSerLeuSerAlaArgHisPro 652
QY 741 CAGCCTTGCAGCACTGGT-----CCACGAGCTGGTGCAGTGCCTCCGTCAGGGCTT 691
Db 653 GlnArgLysArgArgGlyGlyProSerGluProThrProGlySerArgProGlnAspAla 672
QY 690 CCAGCTGGCCCTGT-----CACAGACACCCCGCAGGAATCTGGCCAGCT 646
Db 673 ThrValHisProAlaCysGlnIlePheProHisTyrThrProSerValAlaTyrProTrp 692
QY 645 GCTCCA-----CCA 637
Db 693 SerProGluAlaHisProLeuIleCysGlyProProGlyLeuAspLysArgLeuLeuPro 712
QY 636 TGTCTCAGGTCCTCGATGCTCTTCTCATCTTGAGAAAAGACGTTTCGAGTCCA--- 580
Db 713 GluThrProGlyPro-----CysTyr-----SerAsnSerGlnProVal 725
QY 579 -----TGGGTGCTCCAGAACTCTCGACGCTGCTCAAACAGAGCGGTAGCCAGCT 526
Db 726 TrpLeuCysLeuThrProArgGlnProLeuGluProHisProProGlyGluGlyPro--- 744
QY 525 TATCATTCAAACTCCGCGCAGATTCTTGAAAGGTGCTTGGGGTTCGCGAGCCATAGA 466
Db 745 -----AlaGluGlyArgProCysProTyrPro-----SerGluTrpSerSer 749
QY 465 GAGAGCGGTGGAACCTGATAATAAGACACACAGATCCTTGGGGTTCGCGAGCCATAGA 406
Db 750 AspThr-----AlaGluGlyArgProCysProTyrPro-----SerGluTrpSerSer 749
QY 405 TGACCTTGGAGTCTCCATTGGAGGTGAGAGGTGAGAGGGCAGAAAGCGGT 361
Db 761 -----HisCysGlnValLeuSerAlaGlnProGly 770
RESULT 45
US-10-196-749-236
; Sequence 236, Application US/10196749
; Publication No. US20060094864A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C340
; CURRENT APPLICATION NUMBER: US/10/196,749
; PRIOR FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
```

```
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 236
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-196-749-236
Alignment Scores:
Pred. No.: 1.38 Length: 284
Score: 50.00 Matches: 48
Percent Similarity: 27.6% Conservative: 11
Best Local Similarity: 22.4% Mismatches: 53
Query Match: 2.0% Indels: 102
DB: Gaps: 9
US-10-768-158-1 (1-2419) x US-10-196-749-236 (1-284)
QY 810 TCTCATTTCATGAGACGGTGAAGATGCTCTTCCACAGCCCAACTCTTCCCGGCGCCACGG 751
Db 3 SerTyrProTyrArgGln---GlyCysProGlyAlaGlyGlnAlaProGly----- 19
QY 750 GCAGGGCTCAGCTTGCGACACTGGTCCACCA---GCTGGTGGCAGTCTCCGTCAGGG 694
Db 20 AlaProProGlySerTyrTyrProGlyProProAsnSerGlyGlyGlnTyrGlySerGly 39
QY 693 CTTCAGCTGGGCTTGTTCACAGACACCCCGCAGGAATCTCCGAGCTCTCCACCATCG 634
Db 40 LeuProProGlyGlyGlyTyrGlyGlyProAlaProGlyGlyProTyrGlyProProAla 59
QY 633 TCACGAGTCCCGATGCATGTCTTCATACTTGAGAAAAAGCAGTTCGAGTCCATCGGT 574
Db 60 GlyGlyGly-----ProTyrGly 65
QY 573 GCTCCAGAACTCCTGCACGTGCTCAAACAGGAGCGGTAGCCAGCTTATCATTTCAA 514
Db 66 HisPro----- 67
QY 513 ACCTCCGCGCAGAACTCTTGAAGGTGCTCGGTAGCTCATGTCCGAGAGAGCGGTGGA 454
Db 68 -----AsnProGlyMet 71
QY 453 ACTGATAATAAGACACACCAAGATCCTTGGGGTTGCGAGCCATATAGATGACCTTGGAGT 394
Db 72 PheProSerGlyThrProGlyGlyProTyrGly----- 82
QY 393 CTCCATTGTGGAGTGCAGAGGGCAGAAAGCGTAGGCGAGGTGGCTCTTGTAGAGCGGG 334
Db 83 -----GlyAlaAla----- 85
QY 333 GAGAGTTCAGTTCCTTGTATGATGTCCAGGCGCGGTGTGGGTACTCCAGGACCGGGAGCT 274
Db 86 -----ProGlyGlyProTyrGlyGlnProProSer--- 96
QY 273 GCTCGTCGATGTTTCATCAAGCCGATCTCATCGGGGTGAGCGCCCTGGG----- 226
Db 97 -----SerTyrGlyAlaGlnProGlyLeuTyrGlyGln 108
QY 225 -----TCACCAAGTAGACCACTCTCTGCGACCAAGCTGG 193
Db 109 GlyGlyAlaProProAsnValAspProGluAlaTyrSerTrp 122
RESULT 46
US-11-101-316-62
; Sequence 62, Application US/11101316
```

Publication No. US20060099657A1
GENERAL INFORMATION:
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Wood, William I.
TITLE OF INVENTION: ANTIBODIES TO A POLYPEPTIDE ENCODED BY A NUCLEIC ACID
TITLE OF INVENTION: UNDEREXPRESSED IN MELANOMA
FILE REFERENCE: P3230R1C17C1
CURRENT APPLICATION NUMBER: US/11/101,316
CURRENT FILING DATE: 2005-04-06
PRIOR APPLICATION NUMBER: 10/063526
PRIOR FILING DATE: 2002-05-03
PRIOR APPLICATION NUMBER: 10/006867
PRIOR FILING DATE: 2001-12-06
PRIOR APPLICATION NUMBER: PCT/US00/23328
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: 09/380137
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: PCT/US99/12252
PRIOR FILING DATE: 1999-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
NUMBER OF SEQ ID NOS: 170
SEQ ID NO 62
LENGTH: 284
TYPE: PRT
ORGANISM: Homo Sapien
US-11-101-316-62

Alignment Scores:
Pred. No.: 1-38 Length: 284
Score: 90.00 Matches: 48
Percent Similarity: 27.6% Conservative: 11
Best Local Similarity: 22.4% Mismatches: 53
Query Match: 2.0% Indels: 102
DB: Gaps: 9

US-10-768-158-1 (1-2419) x US-11-101-316-62 (1-284)

Qy	810	TCATCATATGAGACGGTGAAGATGCTCTCCACAGCCCAACTCTTCCCGGCCACGG	751
Db	3	SerTy:ProTy:ArgGln---GlyCysProGlyAlaAlaGlyGlnAlaProGly-----	19
Qy	750	GCAGGCGCTCAGGTTCAGCACTGTGTCCACCA--GCTGTGGCAGTGCTCCGTCCAGG	694
Db	20	AlaProProGlySerTyTyTyProGlyProProAsnSerGlyGlyGlnTyTyGlySerGly	39
Qy	693	CTTCCAGCTGGCGCTGTCTACAGACACACCCAGGAATCTGGCCAGCTCTCCACATCG	634
Db	40	LeuProProGlyGlyTyTyTyGlyGlyProAlaProGlyGlyProTyTyTyProAla	59
Qy	633	TCACCAAGTCCCGATGATGCTTTCATCTTGAGAAAAAGCAGTTCGAGTCCGCGT	574
Db	60	GlyGlyGly-----ProTyTyGly	65
Qy	573	GCTCCCAAGAACTCTGCACGTGCTCAAAACAGAGCGGTAGCCAGCTTATCATTTAA	514
Db	66	HisPro-----	67
Qy	513	ACCTCCGGCAGAATCTTGAAGAGTGCCTCGGTAGCTCATGTTCGACAGAGCGGTGA	454
Db	68	-----AsnProGlyMet	71
Qy	453	ACTGATAAAGACACACACAGATCCTTGGGGTTCGAGCCATATAGATGACCTTGGAGT	394
Db	72	PheProSerGlyThrProGlyGlyProTyTyGly-----	82
Qy	393	CTCATTGTGGAGTTCAGAGCGGAGAAAGCGGTAGGCGAGGTGCTCTTGTAGAGCGGG	334
Db	83	-----GlyAlaAla	85

Qy 333 GAGAGTCACTTCTTGATGATCCAGCCCGGCTGTGGTACTCCAGACCGGAGCT 274
Db 86 -----ProGlyGlyProTyTyGlyGlnProProSer--- 96
Qy 273 GCTCGTCGATGTTTCATCAAGCCGATCTCATCGGGGTTCAGCGCCCTGGC----- 236
Db 97 -----SerTyTyGlyAlaGlnProGlyLeuTyTyGlyGln 108
Qy 225 -----TCACCAAGTAGACACCTCTCTCCAGCAAGCTGG 193
Db 109 GlyGlyAlaProProAsnValAspProGluAlaTySerTyr 122

RESULT 47
US-11-312-958-14
Sequence 14, Application US/11312958
Publication No. US20060100152A1
GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals, Inc.
APPLICANT: Rosenfeld, Julie Beth
TITLE OF INVENTION: METHODS AND COMPOSITIONS IN TREATING
TITLE OF INVENTION: PAIN AND PAINFUL DISORDERS USING 9949, 14230, 760, 62553,
TITLE OF INVENTION: 12216, 17719, 41897, 47174, 33408, 10002, 16209, 314, 636,
TITLE OF INVENTION: 27410, 33260, 619, 15985, 69112, 2158, 224, 615, 44373,
TITLE OF INVENTION: 95431, 22245, 2387, 16658, 55054, 16314, 1613, 1675, 9569 OR
TITLE OF INVENTION: 13424 MOLECULES
FILE REFERENCE: MEI02-027PIRNOHNM
CURRENT APPLICATION NUMBER: US/11/312,958
CURRENT FILING DATE: 2005-12-20
PRIOR APPLICATION NUMBER: US/10/369,022
PRIOR FILING DATE: 2003-02-19
PRIOR APPLICATION NUMBER: US 60/360,495
PRIOR FILING DATE: 2002-02-28
PRIOR APPLICATION NUMBER: US 60/370,121
PRIOR FILING DATE: 2002-04-04
PRIOR APPLICATION NUMBER: US 60/373,010
PRIOR FILING DATE: 2002-04-16
PRIOR APPLICATION NUMBER: US 60/373,908
PRIOR FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: US 60/377,717
PRIOR FILING DATE: 2002-05-03
PRIOR APPLICATION NUMBER: US 60/379,949
PRIOR FILING DATE: 2002-05-13
PRIOR APPLICATION NUMBER: US 60/382,409
PRIOR FILING DATE: 2002-05-21
PRIOR APPLICATION NUMBER: US 60/385,280
PRIOR FILING DATE: 2002-06-03
PRIOR APPLICATION NUMBER: US 60/386,879
PRIOR FILING DATE: 2002-06-06
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 64
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14
LENGTH: 367
TYPE: PRT
ORGANISM: Homo sapiens
US-11-312-958-14

Alignment Scores:
Pred. No.: 1-53 Length: 367
Score: 90.00 Matches: 70
Percent Similarity: 35.5% Conservative: 36
Best Local Similarity: 23.4% Mismatches: 95
Query Match: 2.0% Indels: 98
DB: Gaps: 15

US-10-768-158-1 (1-2419) x US-11-312-958-14 (1-367)

Qy	34	AGGCGGAGACCCCGACACCCCGGGGAGTTCAGGA-----GCAAGTACTTCAGT	84
Db	8	ArgAlaGlyProProGlnProArgAlaArgAlaGlnLeuPheAlaPheThrLeuSer	27
Qy	85	-----TCCATGGCGTGGCTTCTGCGCGGGA	120

```
Db 28 LeuSerCysThrTyrLeuCysTyrSerPheLeuCysCysCys-----AspAspLeuGly 45
Qy 121 AGATGGAGGAGATCGCCAACTCCCGGTGC-----GGCCCGAGGACGTGTGGA----- 168
Db 46 ArgSerArgLeuLeuGlyAlaProArgCysLeuArgGlyProSerAlaGlyGlnLys 65
Qy 169 -----TCGTCACTACCCCAAGT----- 186
Db 66 LeuLeuGlnLysSerArgProCysAspProSerGlyProThrProSerGluProSerAla 85
Qy 187 -----CCGGCACCAGCTTGTGCAGGAGGTGCTACTTGGTGGAGCCAGGCGGTG 237
Db 86 ProSerAlaProAlaAlaAla-ValProAlaProArgLeuSerGly-SerAsnHisSerG 105
Qy 238 ACCCGATGAGATCGGCTTGATGAACATCGACGACGCTC----- 278
Db 105 lySerProLysLeuGlyThrLysArgLeuProGlnAlaLeuIleValGlyValLysLysG 125
Qy 279 -----CCGGTCTCGAGTAC-----C 294
Db 125 lyGlyThrArgAlaValLeuGluPheIleArgValHisProAspValArgAlaLeuGlyT 145
Qy 295 CACAGCG-----GGCCTGGACATCATCAAGNACTGA 327
Db 145 hrGluProHisPhePheAspArgAsnTyrGlyArgGlyLeuAspTrpTyrArgSerLeu- 164
Qy 328 CTTCTCCCGGCTCATCAAGACCACTCCCTACCGCTTCTGCCCTCTGACCTC----- 383
Db 165 --MetProArgThrLeuGluSerGlnIleThrLeuGluLysThrProSerTyrPheValT 184
Qy 384 -----CACAAATGGAGACTCCAAAGTCTATCTATATGG 414
Db 184 hrGlnGluAlaProArgArgIlePheAsnMetSerArgAspThrLysLeuIleValValV 204
Qy 415 CTCCGAACCCCAAGATCTGTGTGTCTTATTATTCAGTTCACCGCTCTCTCGGGACCA 474
Db 204 alArgAsnProValThrArgAlaIleSerAspTyr-----ThrGlnThrL 219
Qy 475 TGAGCTACCGA-----GGCACCTTTCAAGAAATCTCGCGGAGGTTTATGAATGATA 525
Db 219 euSerLysLysProAspIleProThrPheGluGlyLeu-----SerPheArgAsnArgT 237
Qy 526 AGCTGGGC-----TACGGCTCCTGGTTTGAGCACG 555
Db 237 hrLeuGlyLeuValAspValSerTrpAsnAlaIleArgIleGlyMetTyrValLeuHisL 257
Qy 556 TGCAGGAGTTCTGGGAGCAGCGCATGGACTCGAAGCTGCTTTTCTCAAGTATGAAGACA 615
Db 257 euGluSerTrpLeuGlnTyrPheProLeuAlaGlnIleHisPheValSerGlyGluArgL 277
Qy 616 TGCATCGGGACCTGGTGACGATGGTGGAGGAGCTGGCCAGATTCTCGGGGGTG 668
Db 277 euIleThrAspProAlaGlyGluMetGlyArgValGlnAspPheLeuGlyIle 294
```

RESULT 48

```
US-10-505-928-438
; Sequence 438, Application US/10505928
; Publication No. US20060089532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 438
; LENGTH: 717
; TYPE: PRT
; ORGANISM: Homo sapiens
```

RESULT 49

```
US-10-505-928-438
Alignment Scores:
Pred. No.: 1.99 Length: 717
Score: 90.00 Matches: 69
Percent Similarity: 32.9% Conservative: 24
Best Local Similarity: 24.4% Mismatches: 89
Query Match: 2.0% Indels: 101
DB: 6 Gaps: 17
```

US-10-768-158-1 (1-2419) x US-10-505-928-438 (1-717)

```
Qy 37 CCGAGACCCACAGCACCCTCGCGGGGAGATTCGAGACAAGTACTTCGAGTTCATCGCGTGC 96
Db 299 ProGlyProGlyProLysGlyAspProGlyIleGln----- 311
Qy 97 GGCTGCCGCCCTTCGCGGGGAGATCGGAGGATCGCAACTTCCCGGT----- 148
Db 312 -----GlyTyrHisGlyArg-LysGlyGluArgGlyMetProGlyMetProGly 327
Qy 149 -----GCGGCCAGCGACGTTGGATCGTCA 174
Db 327 yLysHisGlyAlaLysGlyAlaProGlyIleAlaValAlaGlyMetLysGlyGluProGly 347
Qy 175 CTTACCCCAAGTCGCGCACACAGCTTGTGCGAGGAGT-----GGTCTACT 219
Db 347 yIleProGlyThrLysGlyGluLysGlyAlaGluGlySerProGlyLeuProGlyLeuLe 367
Qy 220 TGGTGAGCCAGGGCGCTGACCCCGATGAGATCGGCTTGATCAACATCGA----- 268
Db 367 uGlyGlnLysGly-----GluLysGlyAspAlaGlyAsnSerIleGlyGly 382
Qy 269 ----CGAGCAGCTCCCGGT---CCTGGAGTACCC---ACAGCCGGCGCTCGACATCATCA 318
Db 382 yGlyArgGlyGluProGlyProProGlyLeuProGlyProProGlyProLysGlyGluAl 402
Qy 319 AGGAACCTGAC-----CTCTCCCGCTCATCAAGAGCCACCTGCCCTTCTGCG 372
Db 402 aGlyValAspGlyGlnValGlyProProGlyGln----- 413
Qy 373 CCTCTGACCTCCACAATGGAGACTCCCAAGGTTCATCTATATGGCTCGCAA----- 421
Db 414 -----ProGlyAspLysGlyGluArgGlyAlaAlaGlyGlnGlyProAspGly 430
Qy 422 -CCCCAAGATCTCGTGGTGTCTTATTATCAGTTCCACCGCTCTCTCGGACCATGAGCT 480
Db 430 yProLysGlySerLysGly-----GluProGlyLysGlyGluMetValAspTy 446
Qy 481 ACCGAGGACCTTTCAGAAGATTCTGCGGAGGTTTATGAATGATAAGCTGGGTACGGCT 540
Db 446 rAsnGlyAsnIleAsnGluAlaLeuGlnGluIle----- 457
Qy 541 CCTGTTTGACACGCTGACGAGGTTCTGGGAGACCCGACATGAGCTCGAACGTTCTTTTC 600
Db 458 -----ArgThrLeuAlaLeuMetGlyProProGlyLeuPro-----Gly 470
Qy 601 TCAAGTATGAAGACATGCATCGGACCTCTGGTGTGACGATGGT-----GGAGCAGCTGGCCA 654
Db 470 yGlnIle-----GlyProProGlyAlaProGlyIleProGlyGlnLysGlyGly 486
Qy 655 GATT-----CCTGGGGGTGTCTCTGTGACAGGCCAGCTGGAAGCCCTGA 699
Db 486 uIleGlyLeuProGlyProProGlyHisAspGlyGluLysGlyProArgGlyLysProGly 506
Qy 700 CGGAGCACTGCCACCACTGGTGGACCAGTCTGCTCAACGCTGAGGCGCTGCCCGTGGGCC 759
Db 506 yAspMetGlyProPro-----GlyProGlnGlyProPr 517
Qy 760 GGGGAAG 766
Db 517 oGlyLys 519
```

```
US-10-505-928-493
; Sequence 493, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 493
; LENGTH: 744
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-493

Alignment Scores:
Pred. No.: 2.22 Length: 744
Score: 89.50 Matches: 75
Percent Similarity: 25.4% Conservative: 11
Best Local Similarity: 22.1% Mismatches: 76
Query Match: 2.0% Indels: 177
DB: 6 Gaps: 17

US-10-768-158-1 (1-2419) x US-10-505-928-493 (1-744)
QY 783 CTTTCCACAGCCCAACTCTCCCGGCCACGG-----GCAGGGCCTCAG 739
DB 403 ProProGlyAlaIleGlyPheProGlyProLysGlyGluGlyIleValGlyProGln 422
QY 738 CGTTGCAGCATGTCACAGCT----- 715
DB 423 -----GlyProProGlyProLysGlyGluProGlyLeuGlnGlyPheProGly 438
QY 714 -----GGTGGCAGTGTCCGTCAGGGCTT-----CCAGCT 685
DB 439 LysProGlyPheLeuGlyGluValGlyProGlyMetArgGlyLeuProGlyProIle 458
QY 684 GGGCTTGTACAGGACA----- 667
DB 459 GlyProLysGlyGluAlaGlyGlnLysGlyValProGlyLeuProGlyValProGlyLeu 478
QY 666 -----CCCCCAGGA 658
DB 479 LeuGlyProLysGlyGluProGlyIleProGlyAaspGlnGlyLeuGlnGlyProProGly 498
QY 657 ATC-----TGGCCAGCTGTCCACCATCGTCACACAGGTCCCGAT 619
DB 499 IleProGlyIleGlyGlyProSerGlyProIleGlyProProGlyIleProGlyPro--- 517
QY 618 GCATGCTTCATCTTGAGAAAAGACAGCTTCAGTCCATGCGGTGCTCCACAGAACTCCT 559
DB 518 -----LysGlyGluProGlyLeuProGlyProProGlyPhePro 530
QY 558 GCACGTGCTCAACACAGGAGCGGTAGCCCGCTTATCATTAATCAAACTCCCGCAGAAAT 499
DB 531 Gly-IleGlyLysProGlyValAlaGly-----LeuHisGlyProPro--- 544
QY 498 CTTGAAAGTGCCTCGGTAGTCACTGTGTCGACAGAGCGGTGGNACTGATAATAAGACA 439
DB 545 -----GlyLysProGlyAlaLeuGlyProGln----- 553
QY 438 CCACCATCTCTGGGGTTGGAGGCATATAGATGACCTTGGAGTCTCCATTGTGGAGGT 379
DB 553 ----- 553
QY 378 CAGAGGGCAGAAAGCGGTAGGCGAGGTGGCTCTTGATGAGGGGGGAGGTGCTTCTCT 319
DB 554 ----GlyGln--ProGlyLeuProGlyPro----- 561
```

```
318 TGATGATGTCAGGCCGCGCTGTGGGTACTCCAGACCGGAGCTGCTCGTGATGTTCA 259
DB 562 -----ProGlyProPro---GlyProProGlyProProAlaVal----- 573
QY 258 TCAAGCCGATCTCATCGGGGTACAGCCCTCGGTCTACCAAGTAGACCACTCTCTGCGAG-- 201
DB 574 -----MetProProThrProProPro-GlnG 582
QY 200 -----CAAGCTGTGCCGACACTTGGGTAGGTGACGATCCACACGTCGCTGG 154
DB 582 yGluTyLeuProAaspMetGlyLeuGlyIleAaspGlyValLysProProHisAlaTyrgl 602
QY 153 GCGCAC-----CGGGAAGTTGGCGATCTCTCCATCTTCCCGCG-----GCAGAA-- 108
DB 602 yAlaLysGlyLysAaspGlyProAlaTyrgluMetProAlaPheThrAlaGluLe 622
QY 107 ----- 106
DB 622 uThrAlaProPheProProValGlyAlaProValLysPheAaspLysLeuLeuTyraAaspG 642
QY 105 GCGGCAGCCGACGCCCATGGAAGTCTGGAAGTACTTGTCTCGAACTCCCC 57
DB 642 yArgGlnAaspTyraAaspProGlnThrGlyIlePheThrCysGluValPro 658

RESULT 50
US-10-505-928-537
; Sequence 537, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 537
; LENGTH: 1193
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-537

Alignment Scores:
Pred. No.: 2.94 Length: 1193
Score: 89.00 Matches: 97
Percent Similarity: 34.3% Conservative: 41
Best Local Similarity: 24.1% Mismatches: 136
Query Match: 2.0% Indels: 128
DB: 6 Gaps: 24

US-10-768-158-1 (1-2419) x US-10-505-928-537 (1-1193)
QY 1125 CCAGGCCAT-----TGGAACTGCAATGT-----GAGACTGTTTGTAA 1087
DB 463 ProCysHisAaspGlyPheSerCysSerValMetProGluThrGluGluValValCysAasp 482
QY 1086 CAGACATGGAGAGCTGCAGCTTCTAAAGGCGAGACAGCTGC----- 1045
DB 483 AaspCysProProGlyValThrGlyAlaArgCysGluLeuCysAlaAaspGlyTyxPheGly 502
QY 1044 -----TTCGGTTGGGAATCATCACTCCCTCCGCTCACGCCCTCTTCCCTTCCCGCG 991
DB 503 AspProPheGly-----GluHisGly--ProValArgProCysGlnPro----- 516
QY 990 CTGTTTCACAGCTGCTCCAGAGTTTGTCCAGAGTTCCTCCAGAAAGATAAATGAATGCATACAGGAC 931
DB 517 -----CysGlnCysAaspAaspAasp----- 522
QY 930 TTTTGGCTAGTAGACTGTCTGGGTATTGTGAGCAT-----GCAGGTTGTTGTTCTGTT 877
DB 523 --ValAaspProSerAlaAaspGlyAaspCysAaspArgLeuThrGlyArgCysLeuLysCys- 541
```



```

QY 876 ATTATAATAAAGTCAACAGCTGAGGTGCACACTTTCCTCCATCTTCTGTTTATACACCAAGT 817
Db      : : : : :
QY 542 : : : : :
Db      : : : : :
QY 816 CAAACTTCTCATTCATGAGAGCGGTGAAGATGTCTTCCACAGCCCAACTCTTCCCGGC 757
Db      : : : : :
QY 555 laGlyTyr-----PheGlyAspProLeuAlaProAsnP 566
QY 756 CCACGGGAGGGCTCAGCGTTGAGCACTCGGTCCACAGCTGTGGCAGTGCCTCGTCA 697
Db      : : : : :
QY 566 roAlaAspLysCysArgAla-CysAsnCysAsnProMet-----GlySerGluProVal 583
QY 696 GGGCTTCCAGT---GGGCCTGTGCACAGGACACCCCGAAGTCCGCCAGCTGCTCCA 640
Db      : : : : :
QY 584 GlyCysArgSerAspGlyThrCys-----ValCysLys 594
QY 639 CCATCGTCACACAGCTCCGAGTCATCTTTCATACTTGAGAAAAAGCACGTCGTGAGTCCA 580
Db      : : : : :
QY 595 ProGlyPheGlyGlyProAsn---CysGluHis----- 604
QY 579 TGCAGTCTCCAGAACTCTCTGCA-----CGT 553
Db      : : : : :
QY 605 --GlyAlaPheSerCysProAlaCysTyrAsnGlnValLysIleGlnMetAspGlnPhe 623
QY 552 GCTCAAAACAGGAGCGGTAGCCCGACTTATCATTCATAAACCT----- 510
Db      : : : : :
QY 624 MetGlnGlnLeuGlnArgMetGluAla-LeuIleSerLysAlaGlnGlyGlyAspGlyVa 643
QY 509 ---CCGGCAGAAATCTTGAAGGTGCTCGGTAGCTCATGGTCCGAGAGCGGTGGA 454
Db      : : : : :
QY 643 lValProAspThrGluLeuGluGlyArg-----MetGlnGlnAlaGluGlnAlaLeuGl 661
QY 453 ACTGATAATAAGACACACACAGATCTCTTGGGGTTCGAGCCATATAGATGACCTTGGAGT 394
Db      : : : : :
QY 661 nAspIleLeuArgAspAlaGlnIleSerGluGlyAlaSerArg-----SerLeuGlyLe 679
QY 393 CTCATGTGTGAGTGTAGAGGCGAAGAGCGGTAGGCGGTGCTCTTGTATGAGCGGG 334
Db      : : : : :
QY 679 uGlnLeuAlaLysValArgSerGlnGluAsnSerTyrGlnSerArgLeuAspAspLeu-- 698
QY 333 GAGAGGTGAGTTCCTCATGATGCCAGGCCGCTGTGGGTACTCCAGGACCGGAGCT 274
Db      : : : : :
QY 699 ----LysMetThrValGluArgValArgAlaLeuGlySerGlnTyrGlnAsnArg----- 715
QY 273 GCTCGTGCATGTTTCATCAAGCCGATCTCATCGGGGTGAGCGCCCTGCTCACCAAGTAGA 214
Db      : : : : :
QY 716 -ValArgAspThrHisArg-----LeuIleThrGlnMetGl 727
QY 213 CCACCTCTGTCAGCAAGCTGTGCGGACTTGGGTAGGTAGTACGATCCA-----CA 163
Db      : : : : :
QY 727 nLeuSerLeuAlaGluSerGluAlaSerLeuGlyAsnThrAsnIleProAlaSerAspHi 747
QY 162 CGTCGCTGGCGGCAC-----CGGGAAGTTGGCGATCTCTCCATCT 121
Db      : : : : :
QY 747 sTyrValGlyProAsnGlyPheLysSerLeuAlaGlnGlnAlaThrArgLeu----- 764
QY 120 TCCCGCGCAGAGGGCGGAGCGCGACCGCATCGAATCTCGAAGTACTGTCTCTCGAA 63
Db      : : : : :
QY 765 -----AlaGluSerHisValGluSerAlaSerAsnMetGluGlnLeuThrArgGlu 781

```

RESULT 51

```

US-10-511-937-2564
; Sequence 2564, Application US/10511937
; Publication No. US2006008836A1
; GENERAL INFORMATION:
; APPLICANT: EXPRESSION DIAGNOSTICS, INC.
; APPLICANT: Wohlgenuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; APPLICANT: Prentice, James
; APPLICANT: Morris, MacDonaid

```

```

; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; FILE REFERENCE: AND MONITORING TRANSPLANT REJECTION
; CURRENT APPLICATION NUMBER: US/10/511,937
; CURRENT FILING DATE: 2004-10-19
; PRIOR APPLICATION NUMBER: PCT/US2003/012946
; PRIOR FILING DATE: 2003-04-24
; PRIOR APPLICATION NUMBER: US 10/131,831
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US 10/325,899
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 3117
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2564
; LENGTH: 1336
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-511-937-2564

```

```

Alignment Scores:
Pred. No.: 3.07 Length: 1336
Score: 89.00 Matches: 65
Percent Similarity: 35.0% Conservative: 28
Best Local Similarity: 24.4% Mismatches: 109
Query Match: 2.0% Indels: 64
DB: Gaps: 13

US-10-768-158-1 (1-2419) x US-10-511-937-2564 (1-1336)

QY 753 CGGGCAGGGCTCAGCGTTGCAGCAGCTGGTCCACAGCTGGTGGCAGTGTCTCCGTACGGG 694
Db      : : : : :
QY 45 LysAlaGlyProGluGlu---AlaLysAlaProValArgGlyAspGluAlaProProAla 63
QY 693 CTTCACAGCTGGCGCTTCTCAGAGCACACCCCA-----GGAATCGGCCAGCTGCT 643
Db      : : : : :
QY 64 ArgValAlaGlyProAlaAlaGlyThrProProCysGlnMetGlyValThrProThr--- 82
QY 642 CCACCATCGTCACAGCTCCCGATGCATGTCTTCATACTTTGAGAAAAAGCACGTTCCAGT 583
Db      : : : : :
QY 83 -----AspLeuThrLeuGlnLeuLeuAlaValArgArgLysSer 95
QY 582 CCATGCGGTGCTGCC-----AGAACTCTGCAGCTGCTCAACCCAGG----- 541
Db      : : : : :
QY 96 ArgLeuArgAspProGlyLeuGlnGlnThrLeuArgGlyGlnLeuArgLeuGluAsn 115
QY 540 ---AGCGTAGCCAGCTTATCATTAACCTCCGCGCAGAAATCTTGAAGGTGCGCTC 484
Db      : : : : :
QY 116 AspSerArgGluMetAlaArg-----ValLeuGlyGluLeuSerAlaArgLeuLeu 132
QY 483 GGTAGCTCATGGTCCGCGAGAGCGGTGGAACCTGATAATAAGACACCCAGATCCTTGG 424
Db      : : : : :
QY 133 Ser-IleHisSerAspGlnAspArgIleValValThrPheLysThrPheGluGluIleTr 152
QY 423 GGTTCGAGCATATAGATGACCTTGGAGTCTCCATTTGTGGAGGTTCAGAGGCGAAGC 364
Db      : : : : :
QY 152 pLysPheSerThrTyrHisAlaLeuGlyPheThr-----HisHisCys 166
QY 363 GGTAGGCGAGTGGCTTTCATGAGCGGGAGGAGGTTCAGTTCCTTGATGATGCCAGGC 304
Db      : : : : :
QY 166 sLeuAlaAsnLeuLeuMetAspGlnAla----- 175
QY 303 CCGCTGTGGGTACTC-----CAGGACCGGAGCTGCTCGTGCATGTTTCATCAAGC 253
Db      : : : : :
QY 176 ----PheTrpLeuLeuLeuProSerGluGluGluThrAlaIleGlnValHis----- 192
QY 252 CGATCTCATCGGGTCAGCGCTTGGCTTCCACCAAGTAGACCACTCTCTGAGCAAGCTGG 193
Db      : : : : :
QY 193 -----ValAspGluAsnAlaLeuArgLeuThrHisGluSerLeuLeuIleGlnGluGl 210
QY 192 TCGCGGACTTGGGTAGGTAGTCACGATCCACAGCTCGCTGGCGCCGACCGGAAGTTGGCGA 133
Db      : : : : :
QY 210 yProPhePheValLeuCysProAspHisValArg-ValMetThrGly----- 226

```

```
QY 132 TCTCCTCCATCTTCCCGGCGAGAGGGCGCGACGCCGACCGCCACTCGAAGTACT 73
Db 227 -----ProArgAspAlaGly-----AsnGlyProGlnA 236
QY 72 TGCTCTCGAACTCCCGCGGGTG---CTGGGGGTCTCGGCCTCGCTCTCGGCCATGCGCG 16
Db 236 laLeuArgGlnAlaSerGlyAlaProGlnGlyAlaAlaProGluThrAspSerSerp 256
QY 15 CGCGCTCGCGCTCG 2
Db 256 roProSerProSer 260

RESULT 52
US-10-505-928-831
; Sequence 831, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 831
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-831

Alignment Scores:
Pred. No.: 3.98 Length: 2026
Score: 88.50 Matches: 68
Percent Similarity: 27.3% Conservative: 13
Best Local Similarity: 22.9% Mismatches: 121
Query Match: 2.0% Indels: 95
DB: 6 Gaps: 13

US-10-768-158-1 (1-2419) x US-10-505-928-831 (1-2026)
QY 786 TGCTCTCCAGCCGCACTCTCCCG-----GGCCCGAGG 751
Db 770 CysProSerAlaCysGlnCysAsnProHisGlySerTyrGlyThrCysAspProAla 789
QY 750 GCAGGGCTCAGCGTTGCAGCACTGTGCCAGCTGTGGTGCAGTGTCTCGTCAGGGCTT 691
Db 790 ThrGly-----GlnCysSerCysArgProGlyValGlyLeuArg----- 803
QY 690 CCAGCTGGGCTTGTCCACAGGACACCCCGAGAACTTGGCCAGCTGTCTCCACCATCGTCA 631
Db 804 -----CysAspArgCysGluProGlyPheTrpAsnPheArgGlyIleValThr 819
QY 630 CCAGTCCCGATGATGTTCTCATCTACTTGAGAAAAGACGCTTCGAGTCCATGC----- 577
Db 820 AspGlyArgSerGlyCys-----ThrProCysSerCys 830
QY 576 -----GGTGCTCCAGAACTCTCGACGTGTCTCAACACGAG-----AGCCGTAGC 532
Db 831 AspProGlnGlyAlaValArgAspAspCysGluGlnMetThrGlyLeuCysSerCysLys 850
QY 531 CCAGCTTATCTATCATAAACCTCCGGCAGAAATCTTGAA-----AGTGCTCGGTAGCTCA 475
Db 851 ProGlyValAlaGlyProLysCysGlyGlnCysProAspGlyArgAlaLeuGlyProAla 870
QY 474 TGGTCCGAGAGCGGTGGAACTGTATATAAGACACACCATCCATCCTTGGGGTTGGCAG 415
Db 871 GlyCysGluAlaAspAlaSerAla-----ProAlaThrCysAla 883
QY 414 CCATATAGATGACCTTGGAGTCTCCATTGTGGAGTTCAGAGGGCAGAAAGGGTAGGCA 355
```

```
Db 884 GluMetArgCysGluPheGlyAlaArgCys-----ValGluGluSerGlySerAla 900
QY 354 GGTGGCTCTTGATGAGCGGGGAGAGGTCAAGTCTCTTGATGATGTCCAGGCGGCTGTG 295
Db 901 His-----CysValCysProMetLeu 907
QY 294 GGTACTCCAGACCGGAGGCTGCTCGTCGATGTTCATCAAGCGCATCTCATCGGGGT--- 238
Db 908 ThrCysProGluAlaAsnAlaThrLysValCysGlySerAspGlyValThrTyrGlyAsn 927
QY 238 ----- 238
Db 928 GluCysGlnLeuLysThrIleAlaCysArgGlnGlyLeuGlnIleSerIleGlnSerLeu 947
QY 237 -----CAGCGCCCTGGCTCACCAGTAGACCACTCTCTCGAGCAAGC---TGGTGC 190
Db 948 GlyProCysGlnGluAlaValAlaProSerThrHisProThrSerAlaSerValThrVal 967
QY 189 CGGACTTGGGGTAGGTGACGATCCACACGTCGCTGGGCGCGCACCGGGAAGT----- 139
Db 968 ThrThrProGlyLeuLeuSerGlnAlaLeuProAlaProGlyAlaLeuProLeu 987
QY 138 -----TGGCGATCTCTCTCCATCTTCCCGCGGAGAGG 106
Db 988 AlaProSerSerThrAlaHisSerGlnThrThrProProProSerSerArgProArgThr 1007
QY 105 GCGGCGAGCGCGCCCATGGAACCTCGAAGTACTTGTCTCGAACTCCCGCG 55
Db 1008 ThrAlaSerValProArgThrThrValTrpProValLeuThrValProPro 1024

RESULT 53
US-10-505-928-805
; Sequence 805, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 805
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-805

Alignment Scores:
Pred. No.: 1.99 Length: 270
Score: 88.00 Matches: 51
Percent Similarity: 31.1% Conservative: 14
Best Local Similarity: 24.4% Mismatches: 87
Query Match: 2.0% Indels: 58
DB: 6 Gaps: 8

US-10-768-158-1 (1-2419) x US-10-505-928-805 (1-270)
QY 699 TCAGGGCTTCAGCTGGGCGCTTGTTCACAGGACACCCCGAGAACTCGCCAGCTGCTCCA 640
Db 18 AsnGlyLeuHisHisHisProAlaHisArgMetGlyMetGlyGlnPheProSerPro-Hi 37
QY 639 C---CATCTGTCACCGAGTCCCGATGTCATGCTTCTC---ATACTTGAGAAAAGCAGCTCG 586
Db 37 sHisHisGlnGlnGlnProGlnHisAlaPheAsnAlaLeuMetGlyGluHisIleHi 57
QY 585 AGTCCATGCG-----GTGCTCCAGAACTCTCTCGACCTGCTGCTCAACAGGAGC 538
Db 57 sTyrGlyAlaGlyAsnMetAsnAlaThrSerGlyIleArgHisAlaMetGlyProGlyTh 77
QY 537 CGTAGCCCGCTTATCATTAACCTCCGGCAGAAATCTTGAAAGGTGCC----- 486
```

```
Db 77 rValAsnGly-----GlyHisProProSerAlaLeuAlaProAlaAlaArgPheAs 94
Qy 485 -----TCGGTAGCTCATGGTCCG-----
Db 94 nAsnSerGlnPheMetGlyProProValAlaSerGlnGlyGlySerLeuProAlaSerMe 114
Qy 467 -----CAGAGAGCGGTGGAACTGATATAAGACAC-----
Db 114 tGlnLeuGlnLysLeuAsnAsnGlnTyPheAsnHisHisProTyProHisAsnHisTy 134
Qy 437 -----CACCAGATCCTTGGGTGGGTGGAGCCATATAGA 406
Db 134 rMetProAspLeuHisProAlaAlaGlyHisGlnMetAsnGlyThrAsnGlnHisPheAr 154
Qy 405 TGACCTTGGAGT-CTCCATTGTGGAGGTGCAGAGGCGCAGAAAGCGGTAGGCGAGGTGGCTC 347
Db 154 gAspCysAsnProLysHisSerGlyGlySerSerThrProGlyGlySerGlyGlySerSe 174
Qy 346 TTGATGAGCGGGGAGAGGTGAGTTCCTTGATGATGTCAGGCGCGGTGGGTGACTCC 287
Db 174 r-----ThrProGlyGlySerGlySerSe 183
Qy 286 AGGACCGGAGCTGCTCGATGTTTCATCAAGCGATCTCATCGGGGTACGCGCCCTGG 227
Db 183 rGlyGlyGlyAlaGlySerSerAsnSerGlyGlyGlySerGlyGlySerGlyGlySerProAl 203
Qy 226 CTCACCAAGTAGACCACTCTCTGCA 202
Db 203 aSerValAlaHisValProAlaAla 211

RESULT 54
US-10-861-934-16
; Sequence 16, Application US/10861934
; Publication No. US20060089491A2
; GENERAL INFORMATION:
; APPLICANT: Nagata, Shigikazu
; Suda, Takashi
; Takahashi, Tomoniro
; Nakamura, Norio
; TITLE OF INVENTION: A Fas Ligand, A Fragment Thereof and DNA
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/861,934
; FILING DATE: 07-Jun-2004
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/683,703
; FILING DATE: 07-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 1110-163P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-205-8000
; TELEFAX: 703-205-8050
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 278 amino acids
; TYPE: amino acid
```

```
;
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-10-861-934-16

Alignment Scores:
Pred. No.: 2.01 Length: 278
Score: 88.00 Matches: 61
Percent Similarity: 34.5% Conservative: 26
Best Local Similarity: 24.2% Mismatches: 72
Query Match: 2.0% Indels: 93
DB: 6 Gaps: 12

US-10-768-158-1 (1-2419) x US-10-861-934-16 (1-278)
Qy 963 GTCCAGCAAGGAATAAAT---GAATGCATACAGGACTTTTGGCTAGTAGACTGTCTGGGT 907
Db 1 MetGlnProValAsnTyProCysProGlnIleTyTrp----- 14
Qy 906 ATTGTGAGCATGCAGGTGTTGTTTCTGTATTATAAATAAAGTCAAAAGTCAGGTGCAC 847
Db 15 -----ValAspSerSerAlaThrSerProTrp 23
Qy 846 ACTTTCCTCATCTCTGTTTATACACCAAGTCAAACTTCTCATTCATGCAGAGCGGTGAAGA 787
Db 24 AlaProProGlySerValPheSer----- 31
Qy 786 TGCTCTTCCACAGCCCAACTCTTCCCGCGCCACGGCAGGGCCTCAGCGTTGCAGCACT 727
Db 32 CysProSer-----GlyProArgGlyProGlyGlnArg----- 43
Qy 726 GGTCACACAGCTGGTGGCAGTGTCCGTCCAGGGCTTCCAGCTGGGCTTGTGCAGGACA 667
Db 44 ArgProProProProProProSerProLeuPro----- 56
Qy 666 CCCCAGGAATCTGGCCAGCTGCTCCACCATCGTCACCAAGTCC----- 624
Db 57 ProProSerGlnProProProProLeuProLeuSerPro-LeuLysLysLysAspAsnIle 76
Qy 623 -----CCGATGCATGCTTCATCTACTTGGAGAAAAGCAGCGTTCGAGTCCATGC 577
Db 76 eGluLeuTrpLeuProValIlePhePheMetValLeuValAlaLeuValGlyMetGlyLe 96
Qy 576 GGTGCTCCAGAACTCTCGACAGTGTCTCAAAACACAGGAGCC----- 537
Db 96 uGlyMetTyrglnLeuPheHisLeuGlnLysGluLeuAlaGluLeuArgGluPheThrAs 116
Qy 536 -----GTAGCCAGCTTATCATTTATAAACCTCCGGCAGA 502
Db 116 nHisSerLeuArgValSerSerPheGluLysGlnIleAlaAsnProSerThrProSerGl 136
Qy 501 ATTCTTGAAGGTGCTCGGTAGTCTCAT-----GGTCCGACAGAGCGGTGGAACCTCAT 448
Db 136 uThrLysLysProArgSerValAlaHisLeuThrGlyAsnProArgSerArgSerIlePr 156
Qy 447 AATAAGACACACACAGATCCTTGGGGTTGGCGACCATATAGATGACCTTGGAGTCTCCAT 388
Db 156 oLeuGluTrpGluAspThrTyrglyThrAla-----LeuIle 168
Qy 387 TGTGAGGTTCAGAGGGCAGAAAGCGGTAGGCGAGTGGCTCTTGTATGAGGCGGG----- 333
Db 168 eSerGlyValLysTyrlsLys-----GlyGlyLeuValIleAsnGluAlaGlyLeuTy 186
Qy 332 -----AGAGGTTCAG 324
Db 186 rPheValTySerLysValTyPheArgGlyGln 197

RESULT 55
US-10-861-934-26
; Sequence 26, Application US/10861934
; Publication No. US20060089491A2
; GENERAL INFORMATION:
; APPLICANT: Nagata, Shigikazu
```

[illegible]

D6		98	SerProAlaProAlaproProLysProAlaSerProGlyProProGlnValGlu	117
		717		nnnnnnn
Q:		717		nnnnnnn

Db

```

Db 118 ValGlyHisArgGlyGlySerProProArgLeuProProGlyValProValIleSer 137
Qy 675 -----CACAGACACCCCGAGGAATCTGGCCAGCTGCTCCACCATCTCCACAGGTCCC 622
Db 138 LeuGlyHisSerArgProProGlyVal-----AlaMetProThrThrGluLeuGly-Th 155
Qy 621 GATGCATGCTTCATACCTTGAGAAAAACGACGTTTCAGGTCCATCGGTGCTCCAGAACT 562
Db 155 rLeuArg-----ProProLeu 160
Qy 561 CCTGCACGTGCTCAAAACACGAG-----GCCGTAGCCAGCTTATCATATCAATAACCTCCGGC 505
Db 160 uLeuGlnLeuSerThrLeuGlyThrAlaProProThrLeuAlaLeuHisTyrHisProHi 180
Qy 504 AGAATCTTTGAAAGTGTCTCGGTAGCTCATGTGTCGCGAGAGCGGTGGAACTGATTAAT 445
Db 180 sProPheLeuAsn-----SerValTyrIleGlyProAlaGlyProPheSer---IlePh 197
Qy 444 AAGACACACCATGATCCTTTGGGTGGTTCGAGCCATATAGATGACCTTGGAGTCTCCATTGT 385
Db 197 eProSerSerArgLeuLysArgArgProSerHisCysGlu-----LeuAspLeuAl 214
Qy 384 GGAG---GTCAGAGGCGCAAGACGGTAGGCGAGGTGCTCTTGAT-----GAGGCGGG 334
Db 214 aGluGlyHisGlnProGlnLysValAlaArgValPheThrAsnSerArgGluArgTr 234
Qy 333 GAGAGGTTCAGTTCCTTGATGATGTCCAGCCCGCTGTGGTACTC----- 288
Db 234 pArgGlnGlnAsnValAsnGlyAlaPheAlaGluLeuArgLysLeuLeuProThrHisPr 254
Qy 287 -CAGACCGGAGCTGCTCGTCGATGTTTCATCAAGCCGATCTCATCGGGTTCAGCGCCT 229
Db 254 oProAspArgLysLeu-----SerLysAsnGluValLeuArgLeuAlaMetLysTyrIl 272
Qy 228 GCCTCACCAAGTAGACCACTCTCGACCAAGCTGGTCCGAGCTTGGGTAGTGACGA 169
Db 272 eGlyPheLeuVal---ArgLeuLeuArgAspGlnAlaAlaLeuAlaAlaGlyProTh 291
Qy 168 TCACACGTCTCGGCGCG-----CACCGGAAGTTGGCGATCT-CCTCC 125
Db 291 rPro-----ProGlyProAspGlyArgProValHisArgValProAspAspGlyProAr 309
Qy 124 ATCTTCCCGCGCGAGAGGCGGCGACCGCA 94
Db 309 gArgGlySerGlyArgArgAlaGluAlaAla 319

```

RESULT 57

```

US-10-505-928-284
; Sequence 284, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 284
; LENGTH: 620
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-284

```

```

Alignment Scores:
Pred. No.: 2.75 Length: 620
Score: 88.00 Matches: 125
Percent Similarity: 32.8% Conservative: 51
Best Local Similarity: 23.3% Mismatches: 192
Query Match: 2.0% Indels: 169

```

```

DB: 6 Gaps: 27
US-10-768-158-1 (1-2419) x US-10-505-928-284 (1-620)
Qy 1702 CTGTGGAATTCAGCTCAAAATTCGAGCGCAACTCGGTCTCTTTGTGTGCGCGACTCTCCACC 1643
Db 55 ValTrpGlyAlaGluGlyIleProAlaProThrCysTrpIleGlyThrAsp----- 71
Qy 1642 CACTGCATGCAAAATGCTCCTGACTGCTAGACTCCACCTGCCAGCTCACCAGGCGAGA 1583
Db 72 -----ProGlySerProSerArg 77
Qy 1582 CTTCCTCTCAGCACATGCTCCGCGGGCT-----CCTCTGCTATGGGCC 1541
Db 78 -----AlaHisGlnProGlnAlaSerAspAlaAsnArgGluProVal-AlaGluAr 94
Qy 1540 TTTGGAGCCGCCCTCGAAGTCTCTGCGCAGAGGACGTCATCTCTGAAGACCTAGTGAA 1481
Db 94 gSerGluProAlaLeuSerGlyLeuProProAlaThrMetGlySerGlyAspLeuLeu-- 113
Qy 1480 CAATTTCTGCTCATTTTCAGGAGGTTCTGAAGAAGTGAACCTCCCTCAGATAAACAAGTG 1421
Db 114 -----LeuSerGlyGluSerGlnValGluLysThrLysLeuSerSerSe 128
Qy 1420 TCCAGATTCCCCCAACAGACCCCAACTC-----ATCCTCTGTTCTCACCACCTA 1373
Db 128 rGluGluPheProGlnThrLeuSerLeuProArgThrThrLysCysSerGlyHisAs 148
Qy 1372 TCTCAGAACCAAGTAGAAGACAGAAATGGCTGT-----GCCTTAGATAATAAACA 1325
Db 148 pAlaAspThrGluAspAspProSerLeuAlaAspLeuProGlnAlaLeuAspLeuSerGl 168
Qy 1324 TACAAAAATAAGATTAACCTGAGATCTTTTACCATCTGGGTGCGGTCTCCTGA 1265
Db 168 nGlnProHisSerSer-----GlyLeuSerCysLeuSerGlnTrpLysSerValLe 185
Qy 1264 TTCCCTTTGGAAATGAACCTTTTATTGTTTACTGACATTTATGTAGATTTCCAGTGAAA 1205
Db 185 uSerPro----- 187
Qy 1204 AGCTCTATAAAATACAAATAATACGCGGTGTAAGAGGAGCAGACATTTCTAGTGCATATTA 1145
Db 188 -----GlySerAlaAlaGlnProSerSerCys-SerIle- 198
Qy 1144 CAGGCTTTATCCTTACGCTCCAGGCCATTTGAACTGCAATGTGGAGACTGTTTGTAAATCA 1085
Db 199 -----SerAlaSerSerThrGlySerSerLeuGln-----G 209
Qy 1084 GACATGGAGAGGCTGCACGTTCTAAAGCGCAGACAGCTTCGTTGTTGGGAATCATCACA 1025
Db 209 lyHisGlnGluArgAlaGluProArgGlySerLeuAlaLysVal---SerSer-Ser 227
Qy 1024 CTCCCTCCGCTCACGCGCTCTTCCCTTCCCGCTGTTTCA-----CACGCTGCT 974
Db 228 LeuGluProValProGlnGluProSerSerValValGlyLeuGlyProArgProGln 247
Qy 973 TCCAGAGTTTGTCCAGCAAGGAATAAATGAATGATCATCAGACAGCTTTGGCTAGTAGACTG 914
Db 248 TrpSerProGlnProValPheSerGlyGlyAspAlaSerGlyLeuGlyArgArgLeu 267
Qy 913 TCTGGGTATTGTGAGCATGCGGTTGTTGTTTCTGTTATTATTAATAAATAAAGTCAACGTG 854
Db 268 SerPheGlnAlaGluTyrTrpAlaCys-Val----- 277
Qy 853 AGGTACACACTTTTCCCATCTTCTGTTTATACACCAAGTCAAACTTCTCATTTCA-----TG 800
Db 278 -----LeuProAspSerLeu---ProProSerProAspArgHisSerProLeuTr 293
Qy 799 GAGACGGTGAAGATGCTCTTCC-----ACAGCCCCAACT 767
Db 293 pAsn-----ProAsnLysGluTyrGluAspLeuLeuAspTyrThrTyrProLe 309
Qy 766 CTTCCCGCGGCCCGGCGGCGCTCAGCGTTGCGACACTGGTGTCCACAGCT----- 715

```

```
Db 309 uArgProGlyProGlnLeu---ProLysHisLeuAspSerArgValProAlaAspProVa 328
Qy 714 -----GGTGGCAGTGTCCGTGAGGGCTTCCAGCTGGGCTTGTGCAC-----AG 671
Db 328 lLeuGlnAspSerGlyValAspLeuAspSerPheSerValSerProAlaSerThrLeuLy 348
Qy 670 GACACCCCGCAGGATCGGCCAGCTGTCTCCACCATCGTCACCAAGTCCCGATGATGTCT 611
Db 348 sSerProThrAsnValSerProAsnCysProAla----- 360
Qy 610 TCATACTTGAAAAAGCAGCTGTCAGTCCATGCGGTGCTCCAGAACTCCTGCACGTGC 551
Db 361 -----GluAlaThrAlaLeuProPheSerGlyProArgGluProSerLeuLy 376
Qy 550 TCAAAACCGAGCGGTAGCCAGCTTATCATATCAATAAACCTCCGCGAGAAATCTTTGAAAG 491
Db 376 sGlnTrpProSerArgValProGlnLysGln-----GlyGlyMetGlyLeu----- 391
Qy 490 GTGCCTCGGTAGCTCATGCTGCGCAGAGAGCGGTGGAACTGATATTAAGACACCAACAGA 431
Db 392 -----AlaSerTrpSerGlnLeuAla-----SerThrPro----- 401
Qy 430 TCCTTGGGTTGCGAGCATATAGATGACCTTGGAGTCTCCATTTGTGGAGTCCAGAGGC 371
Db 402 -----ArgAlaProGlySerArgAspAlaArgTrpGluArgAr 414
Qy 370 AGAAAGCGGTAGCGAGTGGCTTTGATGAGCGGGAGAGGTGAGTTCCTTGCATGATG 311
Db 414 gGluProAlaLeuArgGly-AlaLysAspArgLeuThrIleGlyLysHisLeuAsp-Met 433
Qy 310 TCCAGGCCCGGCTGTG-----GGTACTCCAGGACGGGAGTGTCTGCTCG 266
Db 434 Gly-SerProGlnLeuArgThrArgAspArgGlyTrpProSerProArgProGluArgG 453
Qy 265 ATGTTTCATCAAGCGATCTCATCGGGTTCAGCGCCT 229
Db 453 uLysArgThrSerGlnSerAlaArg-----ArgPro 463
```

RESULT 58

```
US-10-511-937-2446
; Sequence 2446, Application US/10511937
; Publication No. US2006008836A1
; GENERAL INFORMATION:
; APPLICANT: EXPRESSION DIAGNOSTICS, INC.
; APPLICANT: Wohlgemuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: LV, Ngoc
; APPLICANT: Prentice, James
; APPLICANT: Morris, MacDonald
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND MONITORING TRANSPLANT REJECTION
; FILE REFERENCE: 50661200104
; CURRENT APPLICATION NUMBER: US/10/511,937
; CURRENT FILING DATE: 2004-10-19
; PRIOR APPLICATION NUMBER: PCT/US2003/012946
; PRIOR FILING DATE: 2003-04-24
; PRIOR APPLICATION NUMBER: US 10/131,831
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US 10/325,899
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 3117
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2446
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-511-937-2446
Alignment Scores:
Pred. No.: 2.41 Length: 345
```

```
Score: 87.50 Matches: 64
Percent Similarity: 30.8% Conservative: 22
Best Local Similarity: 22.9% Mismatches: 86
Query Match: 2.0% Indels: 107
DB: 6 Gaps: 15
US-10-768-158-1 (1-2419) x US-10-511-937-2446 (1-345)
Qy 762 CCGCGCCCGCAGGGCAGGGCTCAGCTTGCGAGCACTGGTCCACCAGCTGGT-----GGC 709
Db 50 ProAlaAlaArgProGlyProArg-----ProProAlaGlyGluLeuGly 64
Qy 708 AGTGCTCCGTGAGGGCTTCAGCTGGGCTTGTGCACAGGACA-----CCCCCAGGA 658
Db 65 Ser-----lIleGlyAspHisGluArgAlaAlaAspPheSerProTyr 78
Qy 657 ATCTGGCCAGCTGCTCCACCATCGTCACAGGTCCCGATGCATGTCTTCTATCTAGAA 598
Db 79 LeuGluProLeuGlyAlaProGlnAlaProAlaProAlaThrAlaThrAspThrPheGlu 98
Qy 597 AAAGCAGCTTCGAGTCCATCGGTCTCCAGAACTCCTGCACGT----- 553
Db 99 AlaAlaProProAlaProAlaProAlaProAlaSerGlyGlnHisAspPheLeu 118
Qy 552 -----GCTCAAAACCGAGGAGCGGTAGCCAGCTTATCATTTCA 517
Db 119 SerAspLeuPheSerAspAspTyrGlyGlyAsnCysLysLysProAlaGluTyrGly 138
Qy 516 TAAACCTCC---GGCAGAAATTTGAAAGGTGCTCGTAGCTCATCGTCCGAGAGAGC 460
Db 139 TyrValSerLeuGlyArg-----LeuGlyAlaAlaLysGlyAlaLeuHis 153
Qy 459 GTTGGAACTGAT-----AATAAGACACCACTGCT----- 427
Db 154 ProGlyCysPheAlaProLeuHisProProProProProProProAlaGluLeu 173
Qy 426 -----TGGGTTGCGAGCCATATAGATGACCTTGGAGTCTCCATTTGGAGGTGAG 376
Db 174 LysAlaGluProGlyPheGluPro-----AlaAspCys---LysArg 186
Qy 375 AGGCGAGAAAGCGGTAGGGCAGGTGGCTCTTGTATGAGCGGGGAGAGGTTCAGTTCTTGA 316
Db 187 LysGluGluAlaGlyAlaProGlyGlyGly----- 196
Qy 315 TGATGTCCAGCCCGCTGTGGTACTCCA----- 286
Db 197 -----AlaGlyMetAlaAlaGlyPheProTyrAlaLeuArgAlaTyrLeuGlyTyrGln 214
Qy 285 -----GGACCGGAGCTGCTCGTCGATGTTTCATCAAGCCGATCTCATCGGGT 238
Db 215 AlaValProSerGlySerSerGlySerLeuSerThrSerSerSerSerSer----- 231
Qy 237 CAGCGCCTTGCTCCACCAAGTAGACCACTCTCTGSCAGCAAGCTGGTCCCGAGTTGGGGT 178
Db 232 ---ProGlyThrProSer----- 237
Qy 177 AGGTGAGATCCACACACGTGCTGGCGCCGAGGAGAGTGGCGCATCTCTCCATCTTCC 118
Db 238 -----ProAlaAspAlaLysAlaProProThrAla 247
Qy 117 CGCGGAGAAAGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCG 61
Db 248 CysTyrAlaGlyAlaAlaProAlaProSerGlnValLysSerLysAlaLysLysThr 266
RESULT 59
US-11-101-316-16
; Sequence 16, Application US/11101316
; Publication No. US20060099657A1
; GENERAL INFORMATION:
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
```


Score: 87.50 Matches: 79
Percent Similarity: 35.9% Conservativity: 47
Best Local Similarity: 22.5% Mismatches: 126
Query Match: 2.0% Indels: 99
DB: 6 Gaps: 16

US-10-768-158-1 (1-2419) x US-10-196-749-440 (1-1115)

```
Qy 1032 TCATCAGACTCCCTCCGCTCAGCGGCTCTTCCTCCCTCCCGCTGTTTCACACGCTGCTT 973
Db 365 SerGlnArgLeuArgLeuSerArgAlaLeuValLeu-----SerMetGly 382
Qy 972 CCAGAGTTTCCAGCAAGAAATAAATGAATGCATACAGGACTTTGGCTAGTAGACTGT 913
Db 383 ProGlu-----AspGluGlyValTyGlnCysMetAlaGluAenGluValGlySerAla 400
Qy 912 CTGGGTATTGTGACATGACAGGTTGTTGTTCTGTTATTATAATAAAGTCAAAAGCTGA 853
Db 401 HisAlaValGlnLeuArg----- 407
Qy 852 GGTCACTTTCCCATCTCTGTTTATACACCAAGTCAAACTTCTCATTCATGAGACGG 793
Db 408 -----ThrSerArgProSerIleThrProArgLeuTrpGlnAsp 420
Qy 792 TGAAGA---TGCTCTTCCACAGCCCAACTCTTCCCGCCACCGGCGAGGCTCAGCGT 736
Db 421 AlaGluLeuAlaThrGlyThrProProValSerProSerLysLeuGlyAsnProGluGln 440
Qy 735 TGCAGCACTGTTCCACAGCTGGTGGCAGTGCTCCGTCAGGCTTCCAGCTGGGCTTGT 676
Db 441 MetLeuArgGlyGlnProAla-----LeuPro----- 449
Qy 675 CACAGGACACCCCGAGGAATCTGGCCAGCTCTCCACCATCGTCACAGGTCCCGATGCA 616
Db 450 ---ArgProProThrSerValGlyProAlaSer---ProLysCysProGly----- 464
Qy 615 TGTCTTCATCTTGAGAAAGACGCTGCGAGTCCATCGCTCGCTCC-----AGA 565
Db 465 -----GluLysGlyGlnGlyAlaProAlaGluAlaProIleLeuSer 479
Qy 564 ACTCTGCAGCTGCTCAACACGAGCGGTAGCCAGCTTATCATTCATAAACCCTCCGCG 505
Db 480 SerProArgThrSerLysThrAspSer-TyrGlu-----LeuValTrpArgProArgHi 497
Qy 504 AGAATCTTGAAGGTGCTCGGTAGCTCATGCTCCGACGAGACGGTGGAACTG----- 450
Db 497 sGlu-----GlySerGlyArgAlaProIleLeuTyTy 508
Qy 449 -ATAATAGACACACCACAG-----ATCCTTGGGCT 421
Db 508 rValValLysHisArgLysGlnValThrAsnSerSerAspTrpThrIleSerGlyIle 528
Qy 420 TCGAGCCATATAGATGACCTGGAGTCTCCATTGTGGAGGTTCAGAGGCGAGAAGCGGT 361
Db 528 eProAlaAsnGlnHisArgLeuThrLeuThrArgLeuAspProGlySerLeuTyGluVa 548
Qy 360 AGGGCAGGTGCTTGTATGAGCGGGGAGAGGTTCAGTTCTTGATGATGATCCAGGCCCG 301
Db 548 lGluMetAlaAlaTyAsnCysAlaGlyGluGlyGln-ThrAlaMetValThr-PheArg 567
Qy 300 GCTGTGGGTACTCCAGGACCGGGA-----GCTGCTCGTCGATGTTTCATCAAGCGATCT 247
Db 568 ThrGlyArgArgProLysProGluLeuMetAlaSerLysGluGlnGlnIleGlnArgAsp 587
Qy 246 CATCGGGGTACGCGCTGGGTTCACCAAGTAGACCACTCTTCGACGACCAAGCTGGTCCGG 187
Db 588 AspProGlyAlaSerProGlnSerSerGlnPro----- 599
Qy 186 ACTTGGGTAGTGCAGATCCACAGCTCGCTGGGCGCGACCGGGAAGTTGGGATCTCTCT 127
Db 600 AspHisGlyArg-----LeuSerPro 606
Qy 126 CCATCTTCCCGCGGAGAAAGGCGGCGACGCGCATGCCATGGAACTCGAAGTACTTGTCTCT 67
```

```
Db 607 ProGluAlaProAspArgProThrIleSerThr-AlaSerGluThrSerValTyValTh 626
Qy 66 CGAACTCCCGCGGCTGCTGGGGT 42
Db 626 rTrpIleProArgGlyAsnGlyGly 634

RESULT 61
US-11-302-678-11
; Sequence 11, Application US/11302678
; Publication No. US2006008881A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Silos-Santiago, Immaculada
; APPLICANT: Venkateswarlu, Karicheti
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: UROLOGICAL DISORDERS USING 1435, 559, 34021, 44099, 25278,
; TITLE OF INVENTION: 641, 260, 55089, 21407, 42032, 46556, 62553, 302, 323,
; TITLE OF INVENTION: 12303, 985, 13237, 13601, 18926, 318, 2058 OR 6351 MOLECULES.
; FILE REFERENCE: MPI02-012P1RNM OMNI
; CURRENT APPLICATION NUMBER: US/11/302,678
; CURRENT FILING DATE: 2005-12-14
; PRIOR APPLICATION NUMBER: US/10/345,680
; PRIOR FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: US 60/349,511
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 60/360,500
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/365,041
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 60/374,063
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/403,468
; PRIOR FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: US 60/414,262
; PRIOR FILING DATE: 2002-09-27
; PRIOR APPLICATION NUMBER: US 60/419,986
; PRIOR FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: US 60/423,809
; PRIOR FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: US 60/429,797
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-11-302-678-11
```

Alignment Scores:
Pred. No.: 2.97 Length: 459
Score: 87.00 Matches: 42
Percent Similarity: 31.8% Conservativity: 12
Best Local Similarity: 24.7% Mismatches: 60
Query Match: 2.0% Indels: 56
DB: 7 Gaps: 8

US-10-768-158-1 (1-2419) x US-11-302-678-11 (1-459)

```
Qy 1002 TCCCTTCCCGCTGTTTTCACACGCTGTTCCAGAGTTTGTCCAGCAAGAAATAATGAA 943
Db 325 SerLeuIleProThrIleIleAsnLeuAlaThrAlaLeuThrSerValGlyValGlySer 344
Qy 942 TGCATACAGGACTTTTGGCTAGTAGACTGCTGGGTATTTGTGAGCATGCAGGTTGTTGTT 883
Db 345 PheLeuCysAsp---Trp-----IleLeu 351
Qy 882 TCTGTTATTATAAATAAGTCABAACGTGAGGTTCACACTTTCCCATCTTCTGTTTATACA 823
Db 352 LeuThrPheMetAsnLysAsnLysValTySerHisLysLysPheAspLysValCysThr 371
Qy 822 CCAAGTCAAACCTTCTCATTCATGAGACGGTGAAGATGTCTCTCCACA----- 775
```


QY 345 TGATGAGCGGGAGAGGTCACTTCTTGATGATGTCAGGCGCGCTGTGGTACTCCA 286
Db |||||
216 -----GlyGlyGlu-:: |||||
QY 285 GGACGGGAGTCTCGTCTGATGTTTCATCAAGCCGATCTCATCGGGTCAGCGCCCTGGC 226
Db |||||
224 AspAspGlyAlaAlaCysArgArgAsnAlaGlyGlnGlyArgArgGlySerGlyGlyAla 243
QY 225 TCACCAAGTACACCTCTCTGCACCAAGCTGGTCCGCGACTTGGGGTAGGTGACGATCC 166
Db |||||
244 ArgGlyAlaAlaGlu-ArgArgArgAlaGlyArg-----G1 256
QY 165 ACACCTCGCTGGCGCGACCGGGAAGTTGGCGCATCTCTCATCTTCGCGGGGAGAAGG 106
Db |||||
256 nHisProLeuGlyProHisArgArg--GlyAlaGlnArgAlaAlaGluArgAlaHisPro 275
QY 105 GCGGAGCGCGACCATGGAAGTCTCGAAGTACTTCTCTCGAATCTCCCGCGGGGCTGG 46
Db |||||
276 AlaAlaAlaValArgValGlyProArgGlnGlyAlaGluProArgGlyHisAspProGly 295
QY 45 GGGTCTCGGCTCGCTCTCCG 25
Db |||||
296 GlyProArgGlnArgAlaPro 302
RESULT 65
US-10-511-937-2549
; Sequence 2549, Application US/10511937
; Publication No. US2006008836A1
; GENERAL INFORMATION:
; APPLICANT: EXPRESSION DIAGNOSTICS, INC.
; APPLICANT: Wohlgemuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Lv, Ngoc
; APPLICANT: Prentice, James
; APPLICANT: Morris, Macdonald
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND MONITORING TRANSPLANT REJECTION
; FILE REFERENCE: 506612000104
; CURRENT APPLICATION NUMBER: US/10/511,937
; PRIOR APPLICATION NUMBER: 2004-10-19
; PRIOR FILING DATE: 2003-04-24
; PRIOR APPLICATION NUMBER: US 10/131,831
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US 10/325,899
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 3117
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2549
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-511-937-2549
Alignment Scores:
Pred. No.: 3 39 Length: 394
Score: 86.00 Matches: 59
Percent Similarity: 32.1% Conservative: 22
Best Local Similarity: 23.4% Mismatches: 89
Query Match: 2.0% Indels: 82
DB: 6 Gaps: 11
US-10-768-158-1 (1-2419) x US-10-511-937-2549 (1-394)
QY 1 GCGACGGCGCGCGCGCATGTCGCGAGAGCGGAGAGCCCGACACCCCGCGGGG 60
Db |||||
186 AlaHisAlaAlaAlaGlyProGlyGluValLeuAlaLysGluProProAlaProArg--- 204
QY 61 AGTTCGAGACGACGACTCTCGAGTTCATGGCGTGGCGCTCTTCTGCGCGCGGA 120
|||

Db 205 -----AlaPro 206
QY 121 AGATGGAGGAGATCGCCAACTTCCGGTGCAGCGCCAGCGAGCTGTGGATCGTCACTACC 180
|||::: |||||
207 ArgAlaThrArgAlaProValAla-----SerProAlaAlaLeuGlySerThrAlaThr 224
QY 181 CCAAGTCCG---GCACGAGCT-----TGCTGACGAGG 210
||| |||||
225 AlaSerProAlaAlaProAlaArgArgGlyLeuGluCysSerGluCysLysLeu 244
QY 211 TGGTCTACTTGGTGAGCCAGCGCGTGCACCCGATGAGATCGGCTTGATGAACATCGACG 270
::: |||||
245 PheSerThr----- 247
QY 271 AGCAGCTCCCGTCTCTGGAGTACCACAGCCGCGCTGGACATCATCAAGAACTGACCT 330
||| |||||
248 -----GluThrSerLeuGlnValHisArgArgIleHisThrGlyGluArgProTyrPro 265
QY 331 CTCCCGCGCTCATCAAGAGCCACCTGCCCTACCGCTTCTGCCCTCTCACCTCCACAATG 390
||| |||||
266 CysPro-----AspCysGlyLysAlaPheArgGlnSerThr--HisLeuL 280
QY 391 GAGACTCCAAGTGCATCTATATGCTCGCAACCCCAAGGATCTGGTGGTGTCTTTATTATC 450
|||::: |||||
280 ysAspHisArgArgLeuHisThrGlyGluArgProPheAlaCysGluValCysGlyLysA 300
QY 451 AGTTCACCGCTCTCTCGGACCATGAGCTACCGAGCACCTTTCAAGATTTCTGCCGGA 510
||| |||||
300 laPheAlaIleSerMetArgLeuAlaGluHis-----ArgA 312
QY 511 GGTATTGAATGATAAGCTGGGCTAC-----GGCTCTGGTTTGAGCAG 555
|||::: |||||
312 rgIleHisThrGlyGluArgProTyrSerCysProAspCysGlyLysSerTyrArgSerP 332
QY 556 TGCAGGAGTCTCTGGGAGCACCGCATGGACTCGAACGCTCTTTTCTCAAGTATGAAGACA 615
|||::: |||||
332 heSerAsnLeuTrpLysHisArgLys-----ThrHisGlnGlnG 345
QY 616 TGCATCGGGACCTGCTGACGATGGTGGAGCAGCTGGCCAG-----TTCTTGGGGG 666
|||::: |||||
345 lnHisGlnAlaAlaValArg-----GlnGlnLeuAlaGluAlaAlaValGlyL 363
QY 667 TGTCTGTGACAAGCCCGACGCTGGAAGCCCTG 698
::: |||||
363 euAlaValMetGluThrAlaValGluAlaLeu 373
RESULT 66
US-11-246-999-48
; Sequence 48, Application US/11246999
; Publication No. US20060099622A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 12 Human Secreted Proteins
; FILE REFERENCE: PF489P2
; CURRENT APPLICATION NUMBER: US/11/246,999
; CURRENT FILING DATE: 2005-10-11
; PRIOR APPLICATION NUMBER: US/09/984,130
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243,792
; PRIOR FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: 09/836,353
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/198,407
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: PCT/US99/25031
; PRIOR FILING DATE: 1999-10-27
; PRIOR APPLICATION NUMBER: 60/105,971
; PRIOR FILING DATE: 1998-10-28
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 48
; LENGTH: 995
; TYPE: PRT

US-11-246-999-48

ORGANISM: Homo sapiens

Alignment Scores:

Pred. No.:	4.86	Length:	995
Score:	86.00	Matches:	89
Percent Similarity:	33.0%	Conservative:	48
Best Local Similarity:	21.4%	Mismatches:	135
Query Match:	2.0%	Indels:	143
DB:	7	Gaps:	19

US-10-768-158-1 (1-2419) x US-11-246-999-48 (1-995)

QY 2298 CCATACAGTACTAAATAGAAAAATATAATGAAATTTTCAGAAAGCAGCCCTCCCTCA 2239
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 646 ProThrLysAlaProArgLysLysThrProThrLysPheProThrLysValProProThr 665
QY 2238 CAGAAACACAGGCAAGGTGCTGCGAGGCCACC----- 2206
Db 666 ProThrLysPheProThrLysValProProThrProThrGlnPheProThrLysValPro 685
QY 2205 -----ATCGATGCGACCTGCCACTCGCTGGCCACGGCTTC----- 2170
Db 686 ProThrLysValProSerLysValProProThrProThrGlnPheProThrLys 705
QY 2169 -----TCCCATTAACAACAAAACACAGCTTCTCTACACC----- 2130
Db 706 ValProProThrProThrLysValSerThrLysValLeuSerThrProThrLysAlaHi 725
QY 2129 -----GTTTCAATACAGCACCAAGAGAGACCCCTTCCACAGCAGGACCCCTCCT 2077
Db 725 sThrLysValSerProThrSer-----ThrLysLeuProSerLysAlaProSer----- 741
QY 2076 GACCACCACTTTGGGTGTTTCCAGCCACATAAATAGACAGAACCCCGTCCCGACGGG 2017
Db 742 -ThrGlnThrMetValProThrLysValHis-----ProThrProThrLys 756
QY 2016 GTTGCTGGTACTGCGCCCGGCTCTCCATCGCTTACGCAATGTGGTGAGAAATTA 1957
Db 756 sLeuProThrLysValProThrLysValSerAsn----- 770
QY 1956 GGCAGTGAAGCAATAAGACCTCACCTTACTCTCCCTCAGATACACGGAGGATGAGGTGG 1897
Db 771 -LysValLeuLeuThrSerProGluLeu----- 779
QY 1896 ACAGACAGCAGCGCGCGTCCCAACCTGATCCGAGTGAAACAGGGCTACACTCGCAAAA 1837
Db 780 -----ProThrLysValProProThrProThrLysLeuProThrLysLeuProThrLys 793
QY 1836 TGGTCTCCACGGCTGAGCGCTCAGGTGCACAGACAAACTGAATGATAAAGACCCGG 1777
Db 793 nAlaProProThr----- 797
QY 1776 GGATTTTAGAATTTTACTATTAAAG-----ATCTGCTTGGCCAGCATTAGTAGGCTA 1726
Db 798 -----SerIleLeuLeuSerProThrProThrLysLeuProThrLysIle-SerL 814
QY 1725 TCACCTTCAGGAAC----- 1712
Db 814 euThrLeuThrSerValProThrLysAsnGlnLeuThrSerAlaLysLeuLeuThrThr 834
QY 1711 -----AAGTAGGCTGTGGAATTCAGTCAATTCGAGCGCAACTCGTCTTGGTG 1657
Db 834 hrLeuProThrLysArgAlaThrLysLeuProSerThrSerThrSerValPro 854
QY 1656 GCGGACTCTCC-----ACCACTGCATCAAAATGCTCGT-----ACTGCAT 1615
Db 854 eAsnThrSerCysIleLeuThrHisValGlnProLysMetLeuProThrGluThrArgV 874
QY 1614 TAGACTCCCACTGCCAGCTCCAGCGGAGACTTCCC-----TCAGACATGCTCCGC 1561
Db 874 alProAsnLysMetProProLysProThrArgIleProThrMetSerMetTyrlleThrL 894

QY 1560 GGGCTCCTCTGCATTGGGCTTTGGAGCGCGCGCTGGAAGTCTCTGCGCAGAGGACGTC 1501
Db 894 yLysPro-----P 897
QY 1500 CATTCTGAAGACCTAGTGAACAAATTTCTGTCATTTTCAGGAAGTCTTGAAGAAGTGAAC 1441
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 897 roLeuLysLysAsnSerAlaLysLysVal-----ThrAspL 909
QY 1440 TCGTCAATAAACAAGTGTCCAGATTCCCCCAACAGACCCCAACTCATCTCTGTCT 1381
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 909 ysArgProThrLysSerProLysThrLysProProLysProProLysSerLysThrS 929
QY 1380 CACCACCTATCTCAGAACCAAGTGAGAACAGATGGCTGTGCTTAGATAATAAACT--- 1324
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 929 erValValAsnGln-ProLysLysLysGluThrLysThrGlyValAsnAsnLysThrLys 948
QY 1323 -----ACAAAATAAAGAGTTAACCTCGAGATCTTCTACCATTCGGGTG----- 1279
Db 949 AsnLeuProProLysAlaLysGluProLysLysLysLysLysThrLysValLysGln 968
QY 1278 -----TGGCTCGCTCCTGATTCCTCCCTGGAAA 1252
Db 969 ProValSerHisTyPheProProGlnLysProLysLys 981
RESULT 67
US-11-312-958-38
; Sequence 38, Application US/11312958
; Publication No. US20060100152A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Rosenfeld, Julie Beth
; TITLE OF INVENTION: METHODS AND COMPOSITIONS IN TREATING
; TITLE OF INVENTION: PAIN AND PAINFUL DISORDERS USING 9949, 14230, 760, 62553,
; TITLE OF INVENTION: 12216, 17719, 41897, 47174, 33408, 10002, 16209, 314, 636,
; TITLE OF INVENTION: 72410, 33260, 619, 15985, 69112, 2158, 224, 615, 44373,
; TITLE OF INVENTION: 95431, 22245, 2387, 16658, 55054, 16314, 1613, 1675, 9569 OR
; TITLE OF INVENTION: 13424 MOLECULES
; FILE REFERENCE: ME102-027P1RNMNIM
; CURRENT APPLICATION NUMBER: US/11/312,958
; PRIOR FILING DATE: 2005-12-20
; PRIOR APPLICATION NUMBER: US/10/369,022
; PRIOR FILING DATE: 2003-02-19
; PRIOR APPLICATION NUMBER: US 60/360,495
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/370,121
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 60/373,010
; PRIOR FILING DATE: 2002-04-16
; PRIOR APPLICATION NUMBER: US 60/373,908
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/377,717
; PRIOR FILING DATE: 2002-05-03
; PRIOR APPLICATION NUMBER: US 60/379,949
; PRIOR FILING DATE: 2002-05-13
; PRIOR APPLICATION NUMBER: US 60/382,409
; PRIOR FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: US 60/385,280
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 60/386,879
; PRIOR FILING DATE: 2002-06-06
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 1570
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-312-958-38
Alignment Scores: 5.81 Length: 1570
Pred. No.: 86.00 Matches: 75
Score:


```

Db 1620 AspThrTrpLeuThrThrAsnLeuPro-----AlaLeuThrValGlySerGlu 1635
Qy 451 AGTTCCACCGCTCTC-----TCCGACCATGAGCTACCGAG 486
Db 1636 SerSerLeuAlaLeuAsgLeuValAsnGlyGlyAspArgCysArgGly---ArgValGlu 1654
Qy 487 GCACCTTTCAGAAATTCGCGGAGGTTTATGAATGATAGCTAGGCTACGGCTCCCTGGT 546
Db 1655 ValLeuTyrArgGlySerTrpGlyThrValCysAspAspSerTrpAspThrAsnAspAla 1674
Qy 547 TTGACGACGTCGAGAGTCTCTGGGAGCACCAGCTGCAAGCTGCG-----594
Db 1675 AsnValValCysArgGlnLeuGly---CysGlyTrpAlaMetSerAlaProGlyAsnAla 1693
Qy 595 -----TTTTTCAAGTATGAAGACATCATCGGACCTGG 630
Db 1694 ArgPheGlyGlnGlySerGlyProIleValLeuAspAspValArgCysSerGlyAsnGlu 1713
Qy 631 TGACGATGCTGAGCAGCTGCCAGATTCTCTGG-----663
Db 1714 SerTyrLeuTrpSerCysProHisGlyGlyTrpLeuThrHisAsnCysGlyHisGlu 1733
Qy 664 -----GGGTGCTCTGTGACAAAGGCCAGCTGGAAGCCCTGACGGAGCAGCTGCC 711
Db 1734 AspAlaGlyValIleCysSerAlaThrGlnIleAsn-----SerThrThr 1748
Qy 712 ACCAGCTGGTGG-----ACGAGTGTGCAACGCTGAGGCCCTGCCCGTGGCGGGGAA 765
Db 1749 ThrAspTrpTrpHisProThrThrThrThrAlaArgProSerSerAsnCysGlyGly 1768
Qy 766 GAGTTG-----GGCTGTGGAAGACATCTTCACCGTCT-----CCATGA 804
Db 1769 PheLeuPheTyrAlaSerGlyThrPheSerSerProSerTyrProAlaTyrTyrProAsn 1788
Qy 805 ATGAGAAGTTTGTACTTGTGTATAAACAGAAAGTGGGAAAGTGTACCTCA-----CGT 858
Db 1789 AsnAlaIysCysValTrpGluLe-----GluValAsnSerGlyTyrArg 1803
Qy 859 TTGACTTTTATTTATATAACAGAAACAAACCTGCTGCTACAAATPACCCAGACAGTC 918
Db 1804 IleAsnLeuGlyPheSerAsnLeuLysLeuGluAlaHis-----HisAsnCysSerPheAsp 1822
Qy 919 TAC 921
Db 1823 Tyr 1823

```

RESULT 69

```

; Sequence 383, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/19178
; CURRENT APPLICATION NUMBER: US/10/505,928
; PRIORITY FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 383
; LENGTH: 720
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-383

```

Alignment Scores:

```

Pred. No.: 5.19 Length: 720
Score: 85.00 Matches: 66
Percent Similarity: 29.2% Conservativity: 25
Best Local Similarity: 21.2% Mismatches: 119
Query Match: 1.9% Indels: 102

```

```

DB: 6 Gaps: 13
US-10-768-158-1 (1-2419) x US-10-505-928-383 (1-720)
Qy 822 CCAAGTCAAATTCCTCATTCATGAGACGCTGAAGATGTCTCTCCACAGCC-----772
Db 186 ProProHisSerAlaHisAlaLeuProArgGluSerCysThrAlaHisAlaAsnSerGln 205
Qy 771 -----CAACTCTTCCCGGC-----CCACGGGAGGCGCTCAGCGT 736
Db 206 AlaAlaThrGlnArgLysProGlyThrLysLeuLeuLeuProArgAlaAlaSerValArg 225
Qy 735 TGCAGCAGCTGCTCCACAGCTGGTGGCAGTCTCCGTCAGGCTTCCAGCTGGGCTTGT 676
Db 226 GlyArgSerIleProGlyAlaAlaGluLysProLysGluIleProAlaSerProSer 245
Qy 675 CACAGGACACCCCGAG-----ATCTGGCCAGCTGTCTCCACCA 637
Db 246 ArgThrLysIleProAlaGluLysGluSerHisArgAspValLeuProAspLysPro---264
Qy 636 TCGTCACAGGT-----CCGATGCTGTCTTCTACTTGAAGAAAGCAGGT 589
Db 265 ---AlaProGlyAlaValAsnValProAlaAlaGlySerHisLeuGlyGlnGlyLysArg 283
Qy 588 TCGAGTCCATGCGGTGCTCCAGAACTCTGCACGCTCAAAACCCAGAGCGGTAGCCCA 529
Db 284 AlaIlePro-----ValProAsnLysLeuGlyLysLysThrLeuLeuLysAlaPro 301
Qy 528 GCTTATCATTAACCTCCGCGAGAAATTTGAAAGGTGCTCGGTAGCTCATGTGTC 469
Db 302 GlySer-----ThrSerAsnLeuAlaArgLysSerSerSerGlyProValTrpSer 318
Qy 468 GCAGAGACGCTGGAACCTGATATAAGACACCACCATCTCTGGGGTTCGACCCATAT 409
Db 319 GlyAlaSerSerAla-----323
Qy 408 AGATGACCTTGGAGTCTCCATTGTGGAGTTCAGAGGCGAGAAAGCGGTAGGCGAGTGGC 349
Db 323 -----323
Qy 348 TCTTGATGAGCGGGGAGGTGAGTTCCTTGATGATGTCAGGCCCGCGGTGTCGGTACT 289
Db 324 -----CysThrSerProAlaValGlyLys 331
Qy 288 CCAGGA-----CCGGGAGCTCTGCTCGATGTTCTATCAAGCCGA 250
Db 332 AlalysSerSerGluPheAlaSerIleProAlaAsnSerSerArgProLeuSerAsnIle 351
Qy 249 TCTCATCGGGTCTAG-----CGCCTGGCTCACCAAGTAGACACCT 208
Db 352 SerLysSerGlyArgMetGlyProAlaMetLeuArgProAlaLeuProAlaGlyProVal 371
Qy 207 CCTGCAGCAAGCTGTGTCGGGACTTGGGTTAGTGTAGCAGTCCACACGCTGCTGGCGCGCA 148
Db 372 GlyAlaSerSerTrpGlnAlaLys-----ArgValAspValSerGluLeuAla 387
Qy 147 CCGGAAAGTTGGCATCTCTCCATCTTCCCGGCGAGAGGGCGGAGCGCAGCCAT 88
Db 388 AlaGluGlnLeuThrAlaProProSerAlaSerProThrGlnProGlnThrProGluGly 407
Qy 87 GGA-----ACTCGAAAGTACTTGTCTCGAACTCCCGGGGTGCTGGGGTCTCGG 37
Db 408 GlyGlyGlnTrpLeuAsnSerSerCysAla-----TrpSerGluSer 421
Qy 36 CCTCGCTCTCCGCAATGCCCGCGCGCTGCGCGTGC 1
Db 422 SerGlnLeuAsnLysThrArgSerIleArgArgArg 433

```

RESULT 70

```

US-10-199-229-9
; Sequence 9, Application US/10199229
; Publication No. US20060099701A1
; GENERAL INFORMATION:

```


DB: US-10-768-158-1 (1-2419) x US-11-252-276-152 (1-1125)

Gape:

16

Qy 840 CCATCTTCGTATTACACCAGTCAAACTTCTCATTTGAGCGGTGAAGATGCTCT 781
||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 674 ProLeuAlaThrThrGlnProAlaIysThrSerLysAlaIysThrGln---Pro 692

Qy 780 TCCACAGGCCAACCTTCCC CGGCC CACGGGCAGGG-----CCTCAG 739
:||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 693 ThrSerLeuProLysGlnProAlaProThrThrSerGlyLeuAsnLysLysPromet 712

Qy 738 GCTTGACACTGGTCCA-----CCAGCTGGTGCAGTGT 703
:|||:||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 713 SerLeuAlaSerGlySerValProAlaLaProHisLysArgProAlaAlaThrAla 732

Qy 702 -----CCGTcaggctccacgtggcgcttgt----- 676
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 733 ThrAlaArgProSerThrLeuProAlaArgAspValLysProLysProIleThrGluAla 752

Qy 675 -----CACAGAcacccccaggaatctgccccagctgcaccacctgcgcacca 628
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 753 LysValAlaGluLysArgThrSerProSerLysProSerSerAlaProAlaLeuLysPro 772

Qy 627 GGT-----CCGAgtcagtcttcatct 604
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 773 GlyProLysThrThrProThrValSerLysAlaThrSerProSerThrLeuValSerThr 792

Qy 603 TGAGAAAAAGCAGTTCCAGTCCATCGGTGTCCAGNACTCTGCACGTGCTCMAACC 544
:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db 793 GlyProSerArgSerProAlaThrThrLeuProLysArgProThrSerIleLysThr 812

Qy 543 AGGAGCGTAGCCCCAGCTTATCATTCAAACAACCTCGGCAGAATCTTTGAAAAGTGCCTC 484
:||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 813 GluGlyLys---ProAlaaspvallysargmetThrAlaLysSerAlaSerAlaaspLeu 831

Qy 483 GGTAAGTCTAAGTCCGAGAGCGGTGGAAGCTGATAATAAGACACCACCAGACTCCTGG 424
||| ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 832 SerArgSerLysThrThrSerAlaSerValLysArgAsnThrThrPro----- 848

Qy 423 GTTTCGNACCATAGATGACCTTGGAGTCTCCATTGTGGAGGTcAGAGGCGAGNAAGC 364
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 848 ----- 848

Qy 363 GGTAGGCGAGGTGGCTTTGATAGCGCGGAGAGGTcAGTTCTTGTATG-----ATGT 310
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 849 -----ThrGly-AlaAlaProProAlaGlyMet-ThrSerThrgValLysPrometS 866

Qy 309 CCAGGCCCGCTGTGGGTACTCCAGGACCGGGAGCTGCTGTCGATGTTTCATCAGCCGA 250
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 866 eRaLaPro-----SerArgSerSerGlyAlaLeuSerValhaSpLysLysProt 892

Qy 249 TCTCA-----TCGGGGTCAGGCGCTGCTCACAAGTAGACCCTCCTCGCAGCA 199
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 882 hrSerThrLysProSerSerSerAlaProArgValiserArgueulaThrThrValSer- 901

Qy 198 AGCTGGTCCGAGTGGGTAGGTAGCACATCCACACGTGCTGGCGGCACCGGGAAGT 139
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 902 ----AlaProasPLeu----LysSerValArgSerLysValGlySerThrGluAsni 918

Qy 138 TGGCGATCTCTCCATCTTCCGCGGCGAGAGGGCGGACCGCAGCCGACCATGGAActCGA 79
:||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 918 le-----LysHisiGInProGlyGlyArgAlaLysValGluLysLu 932

Qy 78 AGTACTTGCTCTCGAACTCCCCCGG-----GTCTGGGGGTCTCGG 37
||| ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 932 ysThrGluAlaAlaThrThrAlaGlyLysProGluproAsnAlaValThrLysAlaAlaG 952

Qy 36 CCTCGCTC---TCGCCcANTGCCCGCGCTCG 8
|||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 952 lySerileAlaSerAlaGlnLysProproAla 962

RESULT 73

```

US-11-252-276-22
; Sequence 22, Application US/11252276
; Publication No. US20060094868A1
; GENERAL INFORMATION:
; APPLICANT: Giuliano, Kenneth A.
; APPLICANT: Bright, Gary
; APPLICANT: Olson, Keith
; APPLICANT: Burroughs-Tencza, Sarah
; TITLE OF INVENTION: A System for Cell Based Screening
; FILE REFERENCE: 97-022-K2-CO
; CURRENT APPLICATION NUMBER: US/11/252,276
; CURRENT FILING DATE: 2005-10-17
; PRIOR APPLICATION NUMBER: 09/713,572
; PRIOR FILING DATE: 2000-11-15
; PRIOR APPLICATION NUMBER: 09/430,656
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: 09/398,965
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 09/031,271
; PRIOR FILING DATE: 1998-02-27
; PRIOR APPLICATION NUMBER: 08/810,983
; PRIOR FILING DATE: 1997-02-27
; PRIOR APPLICATION NUMBER: 60/136,078
; PRIOR FILING DATE: 1999-05-26
; PRIOR APPLICATION NUMBER: 60/106,308
; PRIOR FILING DATE: 1998-10-30
; NUMBER OF SEQ ID NOS: 168
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 1610
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: EYFP-DEVD-MAP4-BFP Construct
US-11-252-276-22

Alignment Scores:
Pred. No.: 7.11 Length: 1610
Score: 85.00 Matches: 75
Percent Similarity: 34.0% Conservative: 38
Best Local Similarity: 22.6% Mismatches: 123
Query Match: 1.9% Indels: 96
DB: 7 Gaps: 16

US-10-768-158-1 (1-2419) x US-11-252-276-22 (1-1610)
Qy 840 CCACTCTTGTTATACACCAAGTCAAACCTTCTCATTCATGTGAG
||| :: |||:::|||||||
Db 920 ProLeuAlaThrThrGlnProAlaLysThrSerThrSerLysAla
|||:::|||||||
Qy 780 TCCACAGCCCAACTCTCCCGCGCCACGGCGAGG-----
|||:::|||||||
Db 939 ThrSerLeuProLysGlnProAlaProThrThrSerGlyClyLeu
|||:::|||||||
Qy 738 CGTTGCACACTGGTGCCA-----CC
|||:::|||||||
Db 959 SerLeuAlaSerGlySerValProAlaAlaProHisLysArgPro
|||:::|||||||
Qy 702 -----CGTCAGGGCTTCAGCTGGCGCTTGT-----
|||:::|||||||
Db 979 ThrAlaArgProSerThrLeuProAlaArgAspValLysProLys
|||:::|||||||
Qy 675 -----CACAGACACCCCGAGGAATCTGGCCAGCTGCTGC
|||:::|||||||
Db 999 LysValAlaGluLysArgThrSerProSerLysProSerSerAla
|||:::|||||||
Qy 627 GGT-----CCCCGA
|||
Db 1019 GlyProLysThrThrProThrValSerLysAlaThrSerProSe
|||
Qy 603 TGAAAAAAGCAGCGTTCGAGTGCATGCGGTGCTCCAGAACCTCC
|||:::|||||||

```

Db 1039 GlyProSerSerArgSerProAlaThrThrLeuProLysArgProThrSerSileLysThr 1058
Qy 543 AGGAGCGTAGCCAGCTTATCATTAACCTCCGGCAGAAATCTTGAAGTGCTC 484
Db 1059 GluGlyLys---ProAlaAspValLysArgMetThrAlaLysSerAlaSerAlaAspLeu 1077
Qy 483 GGTAGCTCGTCCGAGAGCGGCTGGAAGCTGATAATAGACACACCAACCACTTGG 424
Db 1078 SerArgSerLysThrThrSerAlaSerSerValLysArgAenThrThrPro----- 1094
Qy 423 GGTTCGAGCCATATAGATGACCTTGGAGTCTCCATTGTGGAGTTCAGAGGCGAGAAGC 364
Db 1094 ----- 1094
Qy 363 GGTAGGCGAGTGCTCTTGATAGCGGGGAGAGTCACTTCCTTGATG-----ATGT 310
Db 1095 -----ThrGly-AlaAlaProAlaGlyMet-ThrSerThrArgValLysProMets 1112
Qy 309 CCAGCGCGGCTGTGGGTACTCCAGGACCGGAGCTGCTCGATGTTTCATCAAGCGGA 250
Db 1112 eAlaPro-----SerArgSerSerGlyAlaLeuSerValAspLysProT 1128
Qy 249 TCTCA-----TCGGGGTCAGCGCCCTGGCTCACCAAGTAGACACCACTCTCGACGA 199
Db 1128 hrSerThrLysProSerSerAlaProArgValSerArgLeuAlaThrThrValSer- 1147
Qy 198 AGCTGGTCCGACTTGGGTAGTGAGCATCCACAGCTCGCTGGCGGACCGGAGT 139
Db 1148 -----AlaProAspLeu-----LysSerValArgSerLysValGlySerThrGluAenI 1164
Qy 138 TGGCGATCTCTCCATCTTCCCGGCGAGAGGCGGCGGAGCGGAGTGAAGTGA 79
Db 1164 le-----LysHisGlnProGlyGlyArgAlaLysValGluLysL 1178
Qy 78 AGTACTTGCTCTCGAAGTCCCGCGG-----GTGCTGGGGGTCTGG 37
Db 1178 yThrGluAlaAlaThrThrAlaGlyLysProGluProAsnAlaValThrLysAlaAlaG 1198
Qy 36 CTGCTC-----TCGCGCATGCGCGCGCTG 8
Db 1198 lySerIleAlaSerAlaGlnLysProProAla 1208

RESULT 74

US-10-196-749-451
; Sequence 451, Application US/10196749
; Publication No. US20060094864A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Deenoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C340
; CURRENT APPLICATION NUMBER: US/10/196,749
; CURRENT FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24

; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 451
; LENGTH: 1743
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-196-749-451
Alignment Scores:
Pred. No.: 7.34 Length: 1743
Score: 85.00 Matches: 101
Percent Similarity: 31.1% Conservative: 26
Best Local Similarity: 24.7% Mismatches: 173
Query Match: 1.9% Indels: 109
DB: Gaps: 20
US-10-768-158-1 (1-2419) x US-10-196-749-451 (1-1743)
Qy 7 GCGAGCGCGGCGGATCGCGGAGAGCGAGCGGAGAGCCCGAGAGCCCGGGGAGTTGCG 66
Db 933 AlathrAlaAlaAla-----GlyThrGly 940
Qy 67 AGAGCAAGT---ACTTCGAGTTCCATGCGTGGCGTGGCGCCCTTCTGCGCGGGAAGA 123
Db 941 GlyAlaGlyCysCysAlaCysThrCysAlaCysThrCysCysAlaGlyAlaAlaAla--- 959
Qy 124 TGGAGGAGATCGCCAACTTCCCGTGGCGGCGGAGAGTGTGATCGTCACTACCCCA 183
Db 960 -----AlaGlyGlyThrGlyGlyAlaThrAlaGlyAlaGlyThrGly 974
Qy 184 AGTCCGACACAGCTTCTGTCGAGGAGTGTCTACTTGGTGAGCGGCGCTGACCCCG 243
Db 975 ThrThrCysAlaThrCysCys-----CysCysAlaGlyAlaThrThr 989
Qy 244 ATGAGATCGGCTTGAACATCGACGAGCAGCTCCCGGCTCTGAGTACCACAGCGG 303
Db 990 ThrCysCysAlaThrThrThrCysThrGlyCysCysThrCysCysThrAlaCysAlaAla 1009
Qy 304 GCCTGGACATCATCAAGGAACGTACCTCTCCCGCCCTCATCAAGAGCCACCTGCCCTACC 363
Db 1010 ThrCysThrGlyGly-----AlaAlaAlaCysCysAlaThrCysCysThr 1024
Qy 364 GCTTCTGCCCTCTGACCTCCACAATGGAGACTCAAGGTCTATATATGCTCGCAACC 423
Db 1025 CysCysCysGlyAlaAlaGlyAla-----ThrGlyGlyGlyCysAlaThr 1039
Qy 424 CCAAGGATCTGGTGGTCTTATTATCAGTTCCACCGCTCTCTGC-----GGACCA 474
Db 1040 -----CysCysAlaAlaAlaThrGlyCysCysThrThrThrGlyAla 1054
Qy 475 TGAGCTACCGAGGCA-----CCTTTCAAGAAATCTGCCCGGA 510
Db 1055 CysAlaAlaAlaAlaAlaThrGlyCysThrGlyAlaThrThrThrThrCysThrGly 1074
Qy 511 GGTATTATGATGATGAAGCTGGGCTACGGCT-----CCTGTTGAGACGCTGCGAGG 561
Db 1075 GlyAlaAlaThrThrGlyCysAlaAlaAlaGlyAlaGlyAlaCysThrCysCys 1094
Qy 562 AGTTCTGGGAGCAGCGCATGGAGCTCGAAGCGCTTTTCTTCTCA----- 603
Db 1095 CysThrGlyCysAlaGlyGlyThrThrThrCysThrAlaAlaAlaGlyCysAlaAlaCys 1114
Qy 604 -----AGTATGAAGACATGTCATCGGACCTGGTGACCATGCTGGAGCAGCTGG 651


```
Qy 181 CCAAGTCCGGACACGACTTGTGCGAGGAGTGCTACTTGGTGAGCGAGCGCTGACC 240
Db |||||
537 ProValProAlaPro-----AlaLeuVal 544
Qy 241 CCGATGAGATCGGCTGTGATGAACATCGAGGACGACTCCCGGTCTCTGGAGTACCCACAGC 300
Db |||||
545 ProValProAla-----ProAlaAlaAlaGlnAlaSerAlaProAlaGlnThr 560
Qy 301 CGGGCTGGACATCATCAAGAACTGACTCTCCCGGCTCATCAAGAGCCACCTGGCCT 360
Db |||||
561 GlnAlaProThrSerAla-----ProAlaValAlaProThrProAlaProThrPro 577
Qy 361 ACCGCTTCTCCCTCTGACCTCCACATGAGACTCCCAAGTCTCATCTATATGG----- 414
Db |||||
578 Thr-----ProAlaValAlaGlnAlaGluValProAlaSerProAlaThrGlyPro 594
Qy 415 -----CTCGCAACCCCAAGGATCTGGTGTGCTTATTATCAGTTCCACCGCT 462
Db |||||
595 GlyProHisArgLeuSerIleProSerLeu----- 604
Qy 463 CTCTCGGACCATGAGCTACCGAGGACCTTTCAAGAAATCT 504
Db |||||
605 ThrCysAenProAspLysThrAspGlyProValPheHisSer 618

RESULT 80
US-10-511-937-2946
; Sequence 2946, Application US/10511937
; Publication No. US2006008836A1
; GENERAL INFORMATION:
; APPLICANT: EXPRESSION DIAGNOSTICS, INC.
; APPLICANT: Wohlgenuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; APPLICANT: Prentice, James
; APPLICANT: Morris, MacDonald
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND MONITORING TRANSPLANT REJECTION
; FILE REFERENCE: 506612000104
; CURRENT APPLICATION NUMBER: US/10/511,937
; PRIOR FILING DATE: 2004-10-19
; PRIOR APPLICATION NUMBER: PCT/US2003/012946
; PRIOR FILING DATE: 2003-04-24
; PRIOR APPLICATION NUMBER: US 10/131,831
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US 10/325,899
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 3117
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2946
; LENGTH: 823
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-511-937-2946

Alignment Scores:
Pred. No.: 6.63 Length: 823
Score: 84.00 Matches: 73
Percent Similarity: 33.3% Conservative: 25
Best Local Similarity: 24.8% Mismatches: 96
Query Match: 1.9% Indels: 100
DB: 6 Gaps: 17

US-10-768-158-1 (1-2419) x US-10-511-937-2946 (1-823)
Qy 8 CGAGCGCGCGCATGGCGAGAGCGAGCGAGACCCCGAGACCCCGAGCGCGGGA----- 61
Db |||||
493 ArgAenArgGluAenGlyGly-----GlyGlyValProGlyValProGlyAlaGlyGly 510
Qy 62 -----GTTTCGAGCAAGTACTT-----CGAGTTCCATGGCGT 94
Db |||||
```

```
Db 511 GlyAlaAlaLysAlaLeuAlaAspGluLysAlaLeuValLeuGlyLysValMetGluAen 530
Qy 95 GCGGCTCGCGCCCTTCTGCGCGCG-----GAAGATGGAGAGATCCGCCACTTCCCGGT 148
Db |||||
531 ValGlyLeuGlyAlaLeuProGlnTyrGlyGluLeuAlaAspLysGlnLysArgGly 550
Qy 149 GCGGCCCGAGCGACTGTGGATCGTCACTCCCAAGTCCGCCACCGAGCTTGTGTCAGGA 208
Db |||||
551 AlaPheLeuLysArg-----AlaAlaGly 558
Qy 209 GGTGGTCTACTTGGTGAGCCAGGCGGTGACCCCGATGAGATCGGTGTGATGAACATCGA 268
Db |||||
559 GlyGlyAspAlaGlyAspAsp-----AspAspAlaGlyGlyCysGly 572
Qy 269 CGAGCAGCTCCCGTCTCTGGAGTACCCACAGCGGGCTGGACATCATCAAGGAACTGAC 328
Db |||||
573 AspAlaGlyAlaGlyAlaValAsnGlyArgGlyGlyPheAlaProGlyThrGlu 592
Qy 329 CTCTCCCGCGCT-----CATCAAGAGCCACCTGCCCTTCTGCGCTCTGACCT 382
Db |||||
593 ProPheProGlyLeuPheProArgLysProAlaProLeuProSerProGlyLeuAenSer 612
Qy 383 CCACAATGGAGACTCCAAGTCTATCTATATGGTCCCAACCCCAAGGATCTGGTGGTGC 442
Db |||||
613 AlaAlaLysArgIleLys----- 618
Qy 443 TTATTATCAGTTCCACCGCTCTCTGCGGACCATGAGCTACGAGGACCTTTCAAGAAAT 502
Db |||||
619 -----ValGluLysAspLeuGluLeuProPro-----AlaAlaLeu 630
Qy 503 CTGCG-----GAGGTTTATGAATGATAAGCTGCGCTACGGCTCTGTTGAGCAGCT 556
Db |||||
631 IleProSerGluAenValTyrSer-----GlnTrpLeuValGlyTyrAla 645
Qy 557 GCAGGA-----GTTCTGGAGACCGCATGAGCTGGAGCTCGAACGTGCT 595
Db |||||
646 AlaSerArgHisPheMetLysAspProPheLeuGlyPheThrAspAlaArgGlnSerPro 665
Qy 596 TTTTCTCAAGTATGAACACATGATCG-----GGACCT 628
Db |||||
666 PheAlaThrSerSerGluHisSerSerGluAenGlySerLeuArgPheSerThrProPro 685
Qy 629 GGTGAC-----GATGGTGAGCAGCTGGCCAGATCTCTGGGGGTGTCTGTGACAAAGC 682
Db |||||
686 GlyAspLeuLeuaspGlyGlyLeuSerGlyArg-----SerGly 698
Qy 683 CCAGCTGGAAGCCCTGACGAGCAGCTGCCACCGAGCTGGTGAGCAAGCTGTGCAACGCTGA 742
Db |||||
699 ThrAlaSerGly-----GlySerThrProHisLeuGlyGlyPro----- 711
Qy 743 GGCCTCGCGCTGGCGCGGAGAGTTGGCTGTGGAAGGA 784
Db |||||
712 -----GlyProGlyArgProSerSerLysGlyGly 721

RESULT 81
US-10-505-928-81
; Sequence 81, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 81
; LENGTH: 623
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-81
```

Alignment Scores:

Pred. No.: 6.55 Length: 623
 Score: 83.50 Matches: 45
 Percent Similarity: 32.7% Conservative: 7
 Best Local Similarity: 28.3% Mismatches: 69
 Query Match: 1.9% Indels: 38
 DB: 6 Gaps: 6

US-10-768-158-1 (1-2419) x US-10-505-928-81 (1-623)

QY 1 GCGAGCGGACGGCGCGGCGGATGGCGGAGAGAGCGGAGAGCCCGGAGAGCCCGGCGG 60
 DB AlaAlaAlaLeuGlyProAlaGlyGluArgAlaGlySerArgProGlnSerProSerGly 317
 QY 61 AGTTCGAGAGCAAGTACTTCGAGTTCCATGGCGTGGCGGCTCCGCGCTTCTGCGCGGGA 120
 DB 318 Asp-----AlaAspAlaArgGlyAspAlaAlaValTyrCysPheHisAlaAlaGly 335
 QY 121 AGATCGAGGAGA-----TCGCCAATTCGCGGTGGCGGCCCGG 159
 DB 336 GluTrpArgGluLeuThrArgLeuProGluGlyAlaProAlaArgGlyCysGlyLeuCys 355
 QY 160 ACGTGTGATCGTCACTACCTACCAAGTCCGGCACCAGCTTGTCTGAGGAGGTGGTCTACT 219
 DB 356 Val-LeuTyrAsnTyrLeu-----PheValAlaGlyGlyValAlaPr 369
 QY 220 TGGTGAAGCGGCGCGTGCACCCCGATGAGTGGCTTGTGATGAACATCGACGAGCTCC 279
 DB 369 oAlaGlyPro-----AspGlyArgAlaArgPr 378
 QY 280 CGGTCTGGAGTACCCAGCGCGGCGCTGGACATCATCAAGNACTGACTCTCCCGCC 339
 DB 378 oSerAspGlnValPheCysTyrAsnProAlaThrAspSerTrpSerAlaValArgProLe 398
 QY 340 TCATCAAGACGACCTGCCTACCGCTTCTGCGCTCTGACCTCCACAAATGGAGACTCCA 399
 DB 398 uArgGlnAlaArgSerGlnLeuArgLeuLeuAla-----LeuAs 411
 QY 400 AGGTCACTATATGCTCGCAACCCAGGATCTGTGTGTGTCTTATTATCAGTT 454
 DB 411 pGlyHisLeuTyrAlaVal-----GlyGlyGluCysLeuLeuSerVal 425

RESULT 82

US-10-523-014-5
 ; Sequence 5, Application US/10523014
 ; Publication No. US20060094101A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lin, Lin-Ling
 ; APPLICANT: Yannoni, Yvonne
 ; TITLE OF INVENTION: MK2 Interacting Proteins
 ; FILE REFERENCE: 08702-0097-00304
 ; CURRENT APPLICATION NUMBER: US/10/523,014
 ; CURRENT FILING DATE: 2005-02-01
 ; PRIOR APPLICATION NUMBER: USN 60/400,044
 ; PRIOR FILING DATE: 2002-08-02
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 5
 ; LENGTH: 433
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-523-014-5

Alignment Scores:

Pred. No.: 6.25 Length: 433
 Score: 83.00 Matches: 58
 Percent Similarity: 28.2% Conservative: 14
 Best Local Similarity: 22.7% Mismatches: 61
 Query Match: 1.9% Indels: 122
 DB: 6 Gaps: 11

US-10-768-158-1 (1-2419) x US-10-523-014-5 (1-433)

QY 714 GGTGGCAGTGTCTCGTCAAGGGCTTCCAGCTGGGCTTGTCAAGAGACACCCCGAGGAATC 655
 DB 5 GlyGlyCysSerProArgAlaProAlaAlaProGlnProArgProProProAlaLeu 24
 QY 654 TGGCAGCTGTCTCCACCATCTGTCCAGGTCCCGATGATGCTCTTCATACTTACAGAAAAA 595
 DB 25 ProProArg-----ProArgAlaProValPro----- 33
 QY 594 GCACGTTTCGAGTCCATGCGGTCTCCC---AGAACTCTCTGCACGTGTCAAAACCCAGGAGC 538
 DB 34 -----AlaSerArgProGlyArgProLeuLeuThrProAlaArgProCysGlyArgMet 51
 QY 537 CGTAGCCAGCTTATCATTAACCTCCCGCAGAAATCTTGAAGGTGCTCGGTAGC 478
 DB 52 Arg----- 52
 QY 477 TCATGGTCCGAGAGCGGTGGAACGTATAAAGACACCCAGATCTCTTGGGTTGC 418
 DB 53 -----ArgGlySerProGlyPro----- 58
 QY 417 GAGCCATATAGATGACCTTGGAGTCTCCATTGTGGAGGTACAGAGGCGAGAAAGCGGTAGG 358
 DB 59 -----ArgLeuGlyGlySerArgGlyGluArgArg 69
 QY 357 GCAGGTGGCTTTCATGATGAGGGGGGAGAGTCACTTCTGATGATGTCAGGCCCGGCT 298
 DB 70 ProAlaGlyArg-----AspProAla 76
 QY 297 GTGGGTACTCAGACCGGAGCTGCTCGATGTTTCATCAAGCCGATCTCATCGGGT 238
 DB 77 ArgValGlyProGlyGlnGlyLeuArg 86
 QY 237 CAGCGCCCTGCTCACCAGTAGACCACTCTCTGCAGCAAGCTGTGCGGACTTGGGT 178
 DB 87 -----ProAlaArgProGlyProAlaAla---TrpThrGluThr-GlyG 100
 QY 177 AGTGAGCATCCAC-----ACGTCGTGGCGCCGACCGGGAAGTTG----- 137
 DB 100 nGly-11eValHisAlaLeuThrAspLeuSerIleProGlyMetThrSerGlyAsnGlyA 120
 QY 136 --GGATCTCTCTCATC----- 122
 DB 120 snSerAlaSerSerIleAlaGlyThrAlaProGlnAsnGlyGluAsnLysProProGlnA 140
 QY 121 -----TTCCCGCGGCGAGAGGCG 103
 DB 140 lAlleValLysProGlnIleLeuThrHisValIleGluGlyPheValIleGlnGlyA 160
 QY 102 GCAGCCGACGCCATGGAATCTCGAAGTACTTGCTCTCGAAC 62
 DB 160 laAspValSerArgTrpAspAlaArgLeuLeuValGlyAsn 173

RESULT 83
 US-11-311-754-31
 ; Sequence 31, Application US/11311754
 ; Publication No. US20060099635A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Vincent Ling
 ; APPLICANT: Kyriaki Dunussi-Joannopoulos
 ; TITLE OF INVENTION: NOVEL GLU5 MOLECULES AND USES THEREFOR
 ; FILE REFERENCE: GNN-007
 ; CURRENT APPLICATION NUMBER: US/11/311,754
 ; CURRENT FILING DATE: 2005-12-19
 ; PRIOR APPLICATION NUMBER: US/09/667,135
 ; PRIOR FILING DATE: 2000-09-21
 ; NUMBER OF SEQ ID NOS: 38
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 31
 ; LENGTH: 558
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:

```
; NAME/KEY:
; LOCATION:
US-11-311-754-31

Alignment Scores:
Pred. No.: 6.9 Length: 558
Score: 83.00 Matches: 35
Percent Similarity: 33.3% Conservative: 5
Best Local Similarity: 29.2% Mismatches: 41
Query Match: 1.9% Indels: 39
DB: 7 Gaps: 6

US-10-768-158-1 (1-2419) x US-11-311-754-31 (1-558)
QY 37 CCGAGACCCCGAGCAGCC-----CGGGGAGTTGAGAGCAAGTACT 78
D 433 ProArgSerProValProSerProAspGlnGlyValGlnGlySerArgHisArg 452
QY 79 TCGAGTTCCATCGCGTGGCGTCCGCCCTTCTGCGCGGGAAGATGGAGGATCGCCA 138
D 453 ProAlaProMetGlyCys-----ProGluTrpValGlnAlaPro 465
QY 139 -----ACTTCCGGTGGCGCCCGAGCGAG----- 162
D 466 AlaProSerProArgGlyValSerArgAlaGlyProGlyThrGlyAlaGlnProProTrp 485
QY 163 -----TGTGATCGTCACTACCCCAAGTCCGGCACCAGCTTCTGCGAGAGTGGTCT 216
D 486 GlyValGlnGlyGlySerArgHisArgArgProAlaProMetGlyCysProGluTrp--- 504
QY 217 ACTTGGTGCAGCGCGCTGACCCGATGAGATCGGCTTGATGAACATCGACGAGCAGC 276
D 505 -----ValGlnAlaProAlaProSerProArgGlyValSer-ArgAlaG 519
QY 277 TCCGGTCTCGAGTACCCACA-----GCCGGCGCTGGACATCATCA 319
D 519 yProGlyThrGlyAlaGlnProLeuTrpGlyValTrpSerGlySerGlyHisArgGln 538

RESULT 84
US-10-505-928-32
; Sequence 32, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 32
; LENGTH: 811
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-32

Alignment Scores:
Pred. No.: 7.99 Length: 811
Score: 83.00 Matches: 74
Percent Similarity: 21.1% Conservative: 28
Best Local Similarity: 22.6% Mismatches: 113
Query Match: 1.9% Indels: 113
DB: 6 Gaps: 15

US-10-768-158-1 (1-2419) x US-10-505-928-32 (1-811)
QY 5 CGGCGACCGCGCGCATGGCGGAGAGCGAGCGCCGAGACCCCGAGCAGTTCGAGTT 64
D 427 ArgArgArgGlnProLeuProThrAlaAlaGlyGluGluMetThrProProAlaGlyLeu 446
QY 65 CGAGAGCAAGTA-----CTTCGAGTTCCATGGCGTGGCGCTGCCGCCCTT 109

Db 447 AlaGluGluLeuProProGlnProGlnGlnGlnGlnGlyArgPheLeuAlaGlyVal 466
QY 110 CTGCGCGGGAAGATGAGGAGATCGCAACTTCCCGGT---GGGGCCGACGCGTGTG 166
D 467 -----AlaTrpAspGlyAlaAlaArgGluLeuValGlyAsnArgSerAlaLeuArgLeu 484
QY 167 GATCGTCAC-----CTACCCCAAGTCCGGCACCAGCTTCTGTCGAGAGGTTGCTA 217
D 485 SerArgArgGlyProGlyLeuGlnGlnProSerProSerValAlaAlaAlaAlaGly--- 503
QY 218 CTGCTGAGCGCGCGCTGACCCCGATGAGATCGGCTTCTGATGAACATCGACGAGCAGCT 277
D 504 -----ProAlaProGlnSerLeuAspLeuHisLysLys----- 514
QY 278 CCGGTCTCTGAGTACCCACA---GCCGGCGCTGGACATCATCAAGAACTGACCTCTCC 334
D 515 ProGlnArgGlyArgProThrArgAlaAspProAlaLeuAlaGluProThrProThrAla 534
QY 335 CCGCTCATCAGAGCCACCTGCCCTACCGCTTCTGCCCTCTGACCTCCACATGAGGA 394
D 535 SerProGlySerAlaPro-----SerProAlaGlyAspProTrpGlnArg 549
QY 395 CTCAAGGTCTATATATGCTCGCAA-----CCCAAGGATCTGTGGT 439
D 550 AlaThrLysHisArgLeuGlyThrGluHisGlnGlnAlaAlaGlnSerAspGlyGly 569
QY 440 GTCTATTATCAGTTCACCGCTCTCTGCGGAC----- 472
D 570 Ala-----GlyLeuProProLeuValSerAspProCysAspPheAsnLysPheIleLeu 587
QY 472 ----- 472
D 588 CysAsnLeuThrValGluAlaValGlyAlaAspSerAlaSerValArgTrpAlaValArg 607
QY 473 ---CATGAGCTACCGAGGACCTTTCAAGAAATTCTGCCGAGGTTTATGAATCATAACT 529
D 608 GluHisArgSerProArgProLeuGlyGly-----Ala 618
QY 530 GGGCTACGGTCTCTGTTTGAAGCAGTGCAGAGAGTCTTGGGAGCAGCCGATGGAGCTCAA 589
D 619 ArgPheArgLeuLeu-PheAspArg----- 626
QY 590 CGTCTTTTCTCAAGTATGAGACATGATCGGACCTGCTGACGATGTTGGAGCAGCT 649
D 627 -----PheGlyGlnGlnProLysPheHisArgPheVal----- 637
QY 650 GCGCAGATTCCTGGGGTGTCTCTGACAAAGCGCCAGCTGGAAGCCCTGACGAGCAGCTG 709
D 638 -----TyrLeuProGluSerSerAspSerAlaThrLeuArgGluLeuArgGlyAspTh 655
QY 710 CCACGAGCTGGT----- 722
D 655 rProTyrLeuValCysValGluGlyValLeuGlyArgValCysProValAlaProAr 675
QY 723 -GACGAGTGTGCAACCGCTGAGGCCCTGCGCTGGCGGCGGAGAGTTGGCTGTGAA 781
D 675 gAspHisCysAlaGlyLeuValThrLeuProGluAlaGlySerArgGlyGlyValAspTy 695

RESULT 85
US-10-505-928-87
; Sequence 87, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT FILING DATE: 2004-08-27
; CURRENT APPLICATION NUMBER: US/10/505,928
```

```
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 87
; LENGTH: 811
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-87

Alignment Scores:
Pred. No.: 7.99      Length: 811
Score: 83.00      Matches: 74
Percent Similarity: 31.1%      Conservative: 28
Best Local Similarity: 22.6%      Mismatches: 113
Query Match: 1.9%      Indels: 113
DB: 6      Gaps: 15

US-10-768-158-1 (1-2419) x US-10-505-928-87 (1-811)
Qy 5 CGGCGACGGCGGCATGGCGAGAGCGAGCGGCGGAGACCCCGAGACCCCGGGGGAGTT 64
Db 427 ArgArgArgGlnProLeuProThrAlaAlaGlyGluGluMetThrProProAlaGlyLeu 446
Qy 65 CGAGAGCAAGTA-----CTTCGAGTTCCATGGCGTGGCGGTCGCGCCCTT 109
Db 447 AlaGluGluLeuProGlnProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 466
Qy 110 CTGCCGCGGGAAGATGGAGGAGATCGCAACTTCCCGT---GCGGCCCGAGCGTGTG 166
Db 467 -----AlaTrpAspGlyAlaAlaArgGluLeuValGlyAsnArgSerAlaLeuArgLeu 484
Qy 167 GATCGTCAC-----CTACCCCAAGTCCGCGACACAGCTTGCTGCGAGGAGTGGTCTA 217
Db 485 SerArgArgGlyProGlyLeuGlnGlnProSerProSerValAlaAlaAlaGly--- 503
Qy 218 CTGGTGTAGCCAGGCGCTGACCCCGATGAGATCGGCTTGATGAACATCGACGACGACT 277
Db 504 -----ProAlaProGlnSerLeuAspLeuHisLysLys----- 514
Qy 278 CCGGCTCTCGAGTACCCACA---GCCGGGCTGGACATCATCAAGGAACTGACCTCTCC 334
Db 515 ProGlnArgGlyArgProThrArgAlaAlaAspProAlaLeuAlaGluProThrProThrAla 534
Qy 335 CCGCTCATCAAGAGCACCTGCCCTACCGCTTCTGCGCTCTGACCTCCACAAATGGAGA 394
Db 535 SerProGlySerAlaPro-----SerProAlaGlyAspProTrpGlnArg 549
Qy 395 CTCGAAGTCTATATATGGCTCGCAA-----CCCCAAGGATCTGGTGT 439
Db 550 AlaThrLysHisArgLeuGlyThrGluHisGlnGluArgAlaAlaGlnSerAspGlyGly 569
Qy 440 GTCTTATATCAGTTCACCGCTCTCTCGGAC----- 472
Db 570 Ala-----GlyLeuProLeuValSerAspProCysAspPheAsnLysPheIleLeu 587
Qy 472 ----- 472
Db 588 CysAsnLeuThrValGluAlaValGlyAlaAspSerAlaSerValArgTrpAlaValArg 607
Qy 473 ---CATGAGTACCGAGCACCTTTCAAGAATTCTGCGGAGGTTTATGAATGATAAGCT 529
Db 608 GluHisArgSerProArgProLeuGlyGly-----Ala 618
Qy 530 GGGTACGGCTCTCGTTTGGACGCTGCAGGATTTCTGGGAGCACCGCATGGACTCGAA 589
Db 619 ArgPheArgLeuLeu-PheAspArg----- 626
Qy 590 CGTGTCTTTTCTCAAGTATGAAGACATGCATCCGGGACCTGTGTGACGATGTGGAGCAGCT 649
Db 627 -----PheGlyGlnGlnProLysPheHisArgPheVal----- 637
Qy 650 GGCCAGATTCTGGGGGTCTCTGTGTGACAAAGGCCCGAGCTGGAAGCCCTGACGGAGCACTG 709

; Sequence 13, Application US/10199229
; Publication No. US2006009701A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Pradeep
; APPLICANT: Wang, Cong-Yi
; TITLE OF INVENTION: ACTIVATORS OF CYCLIN-DEPENDENT KINASES
; FILE REFERENCE: 5853-210
; CURRENT APPLICATION NUMBER: US/10/199,229
; CURRENT FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 951
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-199-229-13

Alignment Scores:
Pred. No.: 8.5      Length: 951
Score: 83.00      Matches: 63
Percent Similarity: 32.4%      Conservative: 21
Best Local Similarity: 24.3%      Mismatches: 108
Query Match: 1.9%      Indels: 67
DB: 6      Gaps: 13

US-10-768-158-1 (1-2419) x US-10-199-229-13 (1-951)
Qy 34 AGGCGGAGAGACCCCGAGACCCCGGGGAGTTCGAGAGCAAGTACTTCGAGTTCCATGCG 93
Db 77 GlnProGlyProProValProAlaAlaProValProAla-----ProSerLeuAla 93
Qy 94 TCGCGCTGCCCGCTTCTGCGCGGGAAGATGGAGG-----AGATCG 135
Db 94 ProGly-----GluAsnGlyThrGlyAspTrpAlaProArgLeuValPheIleGlu 111
Qy 136 CCAACTTCCGGTGGCGGCGCCAGCGACGTGGATCGCTACCTACCCCAAGTCCGCGACCA 195
Db 112 ProProGly---AlaGlyGlyAlaAlaProSerAlaValProThrArgProProGlyPro 130
Qy 196 GCTTGTCTGAGGAGGTGCTCTACTTGTGTGAGCC-----AGGGCGCTGACCCCGATGAGA 249
Db 131 GlnArgCysArgGluGlnSerAspTrpAlaSerAspValGluValLeuGlyProLeuArg 150
Qy 250 TCGGCTTGTATGAAACATCGACGAGC----- 273
Db 151 ProGlyGlyValAlaGlySerAlaLeuValGlnValArgValArgGluLeuArgLysGly 170
Qy 274 -----ACCTCCCGTCTCTGGAGT 291
Db 171 GluAlaGluArgGlyGlyAlaGlyGlyGlyLysLeuPheSerLeuCysAlaTrpAsp 190
Qy 292 ACCCACAGCGCGGCTGGACATCATCAAGGAAGTCTCTCCCGCTCATCAAGAGCC 351
Db 191 -----GlyArgAlaTrp-HisHisGlyAlaAlaGlyGlyPheLeuLeuArgValAr 208
```

Qy 352 ACCTGCCCTA-----CGCTTTCTGCGCTGACCTCCACAATGG---AGACTCCAAGG 402
Db |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
208 sProArgLeuTyrglyProglyGlyAspLeuLeuProProAlaTrpLeuArgAlaLeuGl 228
Qy 403 TCATCTTA-----TATGGCTGCAACCCCAAGGATCTGGTGCTTATTATC 450
Db |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
228 yAlaLeuLeuLeuAlaLeuSerAlaLeuPheSerGlyLeuArgLeuSerLeuSe 248
Qy 451 AGTTCACCGCTC-----TCTGGGACCATGAGCT 480
Db |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
248 rLeuAspProValGluLeuArgValLeuArgAsnSerGlySerAlaAlaGluGlnGluGl 268
Qy 481 ACCGAGCACCTTCAAGAATTCTCCCGGAGTTTATGAATGATAAGCTGGCTACGGCT 540
Db |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
268 nAlaArgValGlnAlaValArgGlyArg-----GlyThrHisLe 282
Qy 541 CCTGGTTTGGACCGTGCAGAGTTCTGGGAGCACCGCATGGATCGAAGCT---GCTTT 597
Db |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
282 uLeuCysThrLeuLeuLeuGlyGlnAlaGlyAlaAsnAlaAlaLeuAlaGlyTrpLeuTy 302
Qy 598 TTCTCAAGTATGAACATGATCGGACCTGGGACCTGGTGAGCATGGTGAGAGCTGGC 652
Db :::: :::: |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
302 rAlaSerLeuProGlyValGlyAspProGlyGluAspSerGlyGluAlaGly 320

RESULT 87

US-10-196-749-518
; Sequence 518, Application US/10196749
; Publication No. US20060094864A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C340
; CURRENT APPLICATION NUMBER: US/10/196,749
; CURRENT FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 518
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-196-749-518

Alignment Scores:
Pred. No.: 9,11 Length: 1137
Score: 83.00 Matches: 170
Percent Similarity: 28.7% Conservative: 84
Best Local Similarity: 19.2% Mismatches: 292
Query Match: 1.9% Indels: 340
DB: 6 Gaps: 38

US-10-768-158-1 (1-2419) x US-10-196-749-518 (1-1137)
Qy 2262 TTTCACAGAAAGCAGCTCCCTCACAGAAACACAGCAAGTGTGTCGCGAGCCCAATG 2203
Db :::: |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
280 TrpAsnArgSerGluGlyTrpSerArgAsnHisArgAlaGlyGlyProGlu----- 296
Qy 2202 CGATGCCACGTGCACCTGCTGCGCCAGCTTCTCCCATTTACAAACAAAACAACAGCT 2143
Db |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
297 -----ArgProSerTrpProArgArgAspHisSerArgAsnSerAlaGlnGly 312
Qy 2142 TCTCTCTACACCCGTTTCAAAATACAGCACCAAGCGAGACCCCTTCCACAGCAGCGAC 2083
Db |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
313 SerPheLeuValSerVal-----AlaGluProGlyThrProAlaIleLeuSerSer 329
Qy 2082 CCTCCTCACCACTTTGGGTGCTTTCCAGCCACATAAATGACAGCAACCCCGTCCC 2023
Db |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
330 ProGluPheGlnAlaSerGly-----ThrSerAsnCysSerLeu 342
Qy 2022 CAGGGTGTGGTGGTACTGCGCCCGACTCTCCATCGCTTACGCAATGTGGGTGAGA 1963
Db :::: |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
343 ValPheTyrglnTyrlLeuSerGlySerGluAlaGlyCysLeuGln---LeuPheLeuGln 361
Qy 1962 AAATTAGGCAGTGAAGCAATAAAGACCTCACCTTTACTCTCCCTCAGATACACGA--- 1909
Db |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
362 ThrLeuGlyProGlyAlaProArgAlaProValLeuLeuArgArgArgGlyGluLeu 381
Qy 1908 GGGATGAGGTGG-----ACAGACACGAGCGCGCGCTGCCAACCTGATCCGAG 1861
Db |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
382 GlyThrAlaTrpValArgAspArgValAspIleGlnSerAlaTyPro----- 397
Qy 1860 TGAACAGGGCTACATCGCAAAATGTCTCTCCACGGCTGAGCGCTCAGGTGCACAGA 1801
Db 397 ----- 397
Qy 1800 ACAAACTGAATGATAAGACCCCGGGATTTTAGAAATTTACTATTAAAGCATCTGTTGCC 1741
Db 397 ----- 397
Qy 1740 AGCATTAGTAGGTATCATCTTACAGAAACAAAGTAAGGCTGTGGAAATTCAGTCAAATTC 1681
Db 398 -----PheGlnIleLeuLeu 402
Qy 1680 CGAGCGCAACTCGGTCTCTTTGGTGGC-----CGACTCTCCACCC 1642
Db |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
403 AlaGlyGlnThrGlyPro--GlyGlyValValGlyLeuAspAspLeuIleLeuSerAspH 422
Qy 1641 ACTGCATGCAAAATGCTCTGACTGCATTAGACTCCACCTGCCAGCTCAGCTCAGCCAGAC 1582
Db |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
422 iCysArgPro-----ValSerGluValSerThrLeuGlnProLeuProGly---- 438
Qy 1581 TTCCCTCAGCACATGCTCCGCGGCTCTCTGCAATTTGGGCTTTGGAGCCGCCGCT--- 1525
Db |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
439 -----ProArgAlaPro-AlaProGlnProLeuProProSerSerArg 452
Qy 1524 -----GGAAGTCTCTGGCAGAGGAGCTCATCTTGAAGACCTTAGTG-----AAC 1480
Db |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
453 LeuGlnAspSerCysLysGlnGlyHisLeuAlaCysGlyAspLeuCysValProGlu 472
Qy 1479 AATTCTGCTCATTTACAGAA----- 1459
Db |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
473 GlnLeuCysAspPheGluGluGlnCysAlaGlyGluAspGluGlnAlaCysGlyThr 492
Qy 1458 -----GGTTCTGAAGAAGTGAACCTCGTCAGCAATAAAC 1426
Db |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
493 ThrAspPheGluSerProGluAlaGlyGlyTrpGluAspAlaSerValGlyArgLeuGln 512


```
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 97
; LENGTH: 1894
; TYPE: PR1
; ORGANISM: Homo Sapien
US-10-196-749-97

Alignment Scores:
Pred. No.:      11.1      Length:      1894
Score:          83.00     Matches:      83
Percent Similarity: 32.3%  Conservative: 29
Best Local Similarity: 23.9% Mismatches:    150
Query Match:     1.9%     Indels:       85
DB:              6       Gaps:        16

US-10-768-158-1 (1-2419) x US-10-196-749-97 (1-1894)

Qy 61 AGTTGAGAGCAAGTACTTCAGTTCATGCGTGC---GGCTCGCGCCCTTCGCGCG 117
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
235 ThrThrAlaThrThrGlyAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaThr 254
Qy 118 GGAAGATGGAGAGATGCCAACTTCCCGGTGCGCGCCAGCAGAGTGTGGATCGTCACT 177
Db :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
255 GlyAlaGlyAlaThrAlaThrThrGlyAlaCysAlaThrAlaGlyAlaGlyThrGly 274
Qy 178 ACCCAAGTCGGCACCAGTGTGTGCGAGAGGTGCTTACTTGGTGGAGCCAGGCGGTG 237
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
275 Thr-----GlyGlyThrCysCysAlaAlaAlaGly-----AlaAlaGlyCys 288
Qy 238 ACCCGATGAGATCGGCTTGTGAACATCGAGCAGCAGCTCCCGTCTCGGAGTACCCAC 297
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
289 ThrCysCysThrAlaAlaAlaGlyCysThrThrGlyCysAlaGlyAlaAlaAlaThr--- 307
Qy 298 AGCCGGGCTGGACATCATCAAGGAATGACCTCTCCCGCTCATCAAGAGCCACCTGC 357
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
308 -----ThrThrThrAlaThrCys 313
Qy 358 CTAACCGCTTCTGCCCTCTGACCTCCACATGAGAGACTCAAGTCACTATATGGCTC 417
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
314 ---CysAlaAlaCysThrThrGlyThrThrGlyGlyAlaAlaAlaGlyCysThrThr 332
Qy 418 GCAACCCCAAGGATCGTGTGTGCTTATTATTCAGTTCACCGCTCTCTGCGGACCATGA 477
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
333 AlaThrThrAlaThrGlyAlaCys-----AlaAlaThrAlaCysCysAlaThrThr 349
Qy 478 GCTACCGAGGACCTTCAAGAAATTCGCCGAGGTTTATGAATGATGAAGTGGGCTACG 537
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
350 ThrThr-----ThrCysAlaThr 355
Qy 538 GCTCTGTTTGAGCAGCTGAGGAGTCTGGAGCACCAGTACGACTCGAAGCTGCTTT 597
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
356 Ala---GlyAlaGlyThrThrGlyThrGlyCysCysThrGlyThrThrCys--- 373
Qy 598 TTCTCAAGTATGAACATCATCGGACCTGGACCTGTGACGATGGTGAGCAGCTGGC----- 652
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
374 ---AlaThrAlaGlyThrCysCysAlaAla-GlyGlyCysGlyGlyAlaGlyAlaThrCy 392
Qy 653 -----CAGATTCTGGGGTGTCTGTGACAAAGGCCAGCTGGAAGCCCTGACGGAGC 705
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
392 sCysThrAlaCysThrGlyGlyCysAlaCysAlaGly---GlyGlyAlaGlyThrGlyG1 411
```

```
Qy 706 ACTGCCACCAAGCTGGTGACCAAGTGTGCAACGCTGAGGCCCTGCGCGTGGCGGGAA 765
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
411 yThrGlyGlyAlaGlyAlaGlyThr-CysThrAlaThrCysThrAlaThrGly-GlyAla 430
Qy 766 GAGTTGGCTGTGGAAGGACATCTTCACCGCTCTCCATGAATGAGAAGTTTCACTTGGTGT 825
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
431 GlyCysGlyCysCysAlaThrThrCysAlaAla-----AlaGlyAlaThr-Gly-- 446
Qy 826 ATAAACAGAAAGATGGGAAAGTGTGACCTCAGTTTGACTTTTATTATAATAACAGAAAC 885
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
447 -----AlaAlaTh 449
Qy 886 AACAACTGCTGCTCACAATACCCAGACAGTCTACTAGCCAAAGTCCGTGTATGCATTC 945
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
449 rThrThrCysAlaThrThrCysAlaCysGlyGlyThrThr----- 462
Qy 946 ATTTATTCCTTGTGGACAAACTCTGGAAGCAGCG---TGTGAACACGCGGGGAAAGGA 1002
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
463 -----GlyCys-GlyThrThrThrAlaAlaThrCysGlyGlyAlaGlyAlaGly- 479
Qy 1003 AGAGCGCGGTGAGCGGAGGAGTGTGATGATTCACCAACGAGCAGCTGTCTGCCTTTA 1062
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
480 --GlyAlaCysThrGlyGlyThrThrGlyCysCysAlaThrGlyGlyCysAlaAlaAla 499
Qy 1063 GAACGTGCAGC 1073
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
499 hrGlyCysThr 502

RESULT 89
US-10-370-959-40
; Sequence 40, Application US/10370959
; Publication No. US20060088907A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Williamson, Mark W.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: MacBeth, Kyle J.
; APPLICANT: Hunter, John J.
; APPLICANT: Rudolph-Owen, Laura A.
; APPLICANT: Bandaru, Rajasekhar
; APPLICANT: Tsai, Fong-Ying
; TITLE OF INVENTION: NOVEL 13237, 18480, 2245, 16228, 7677,
; TITLE OF INVENTION: 26320, 46619, 33166, 18836, 48867, 21617, 55562, 39228,
; TITLE OF INVENTION: 62088, 46745, 23155, 21657, 42755, 32229, 46863 AND
; TITLE OF INVENTION: 32252 MOLECULES AND USES THEREFOR
; FILE REFERENCE: MPI03-0170MIM
; CURRENT APPLICATION NUMBER: US/10/370,959
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: US 09/910,150
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/219,028
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: US 10/251,507
; PRIOR FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US 09/715,479
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: US 60/218,053
; PRIOR FILING DATE: 2000-07-13
; PRIOR APPLICATION NUMBER: US 09/644,929
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: US 60/212,439
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 09/892,870
; PRIOR FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: US 60/214,174
; PRIOR FILING DATE: 2000-06-26
; PRIOR APPLICATION NUMBER: US 09/775,117
; PRIOR FILING DATE: 2001-02-01
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 156
; SOFTWARE: FastSeq for Windows Version 4.0
```



```
; SEQ ID NO 40
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-370-959-40

Alignment Scores:
Pred. No.:      8.14      Length:      665
Score:          82.50     Matches:      60
Percent Similarity: 32.6%   Conservative: 31
Best Local Similarity: 23.8% Mismatches: 104
Query Match:     1.9%     Indels:      75
DB:              6       Gaps:       13

US-10-768-158-1 (1-2419) x US-10-370-959-40 (1-665)
Qy 789 AGATGCTCTCCACAGCCAACTCTCCCGCCCGCCAGCGGCGCT-----CAGCGT 736
Db 19 GlnCysProValCysGlnGlnMetMetProAlaAlaHisIleAsnSerHisLeuAspArg 38
Qy 735 TGCAGCACTGGTCCACAGCTGGTGGCAGTGTCCGTCAGGCGCTTCAGTGGCGCTTGT 676
Db 39 CysLeuLeuHisProAlaGlyHisAlaGlu-----ProAlaAlaGlySer 54
Qy 675 CACAGGACACCCCGAGGAATCGGCCAGCTGTCCACCATCGTCCACAGCTCCCGATGCA 616
Db 55 HisArgAlaGlyGlu-----ArgAlaLysGlyProSerProGly----- 68
Qy 615 TGTCTTCATCTTGAGAAAGACAGT-----TCAGATCCATCGGTGCTCC 568
Db 69 -----AlaLysArgArgLeuSerGluSerSerAlaLeuLysGlnPro 83
Qy 567 AGAATCTCTGCAGCTGCTCAACACAGAGCCGCTAGCCCGCTATCATTCATAAACCTCC 508
Db 84 AlaThrProThrAlaAlaGluSerSerGlu----- 93
Qy 507 GGCAGAACTCTTGAAGGTGCTCGTAGCTCATGTCGGCAGAGCGGTGGA---ACT 451
Db 94 -----GlyGluGlyGluGlyAspAspGlyGlyGluThr 105
Qy 450 GATAATAAG-----ACACACAGATCCTTGGGGTGGCGACCATATAGATGA 403
Db 106 GluSerArgGluSerTyrAspAlaProProThrProSerGly-AlaArgLeuIleProAs 125
Qy 402 CTTGGAGTCTCCATTGTGGAGTTCAGAGGCAGAAACGGTAGGGCAGGTGGCTTTGA 343
Db 125 pPheProValAlaArgSerSerSerProGlyArgLysGlySerGlyLysArgProAlaAl 145
Qy 342 TGAGCGGGGAGAGT-----CAGTTCCTTGATGATGTCAGGCCCGGCTGTG 295
Db 145 aAlaAlaAlaAlaGlySerAlaSerProArgSerTrpAspGluAlaGluAlaGln----- 163
Qy 294 GGTACTCCAGACCGGAGCTGCTCGTGTGATGTTTCAT-----CAAGCGCATCTCAT 244
Db 164 -----GluGluGluGluAlaValAlaGlyAspGlyAspGlyAspAlaAspAlaAs 181
Qy 243 CGGGGTACAGCCCTCGGTCCACCAAGTAGACACCTCTCGCAAGCTGTGTC---GG 187
Db 181 pGlyGluAspAspProGlyHisTyrAspAlaAspAlaAlaGluAlaAlaThrAlaPheG 201
Qy 186 ACTTGGGTGAGTACGATCCACAGTC---GCTGGCGCCGACCGGGAGTTGGCATCT 130
Db 201 yAlaSerGlyGlyArgProHisProArgAlaLeuAlaAlaGluGluIleArgGlnMe 221
Qy 129 CTTCCAT---CTTCCCGCGGAGAGGCGGCGAGCCGCGCCATG----- 87
Db 221 tLeuGlnGlyLysProLeuAlaAspThrMetArgProAspThrLeuGlnAspTyrPheG 241
Qy 86 -----GAACT 82
Db 241 yGlnSerLysAlaValGlyGlnAspThrLeuLeuArgSerLeuLeuGluThrAsnGluI 261
Qy 81 CGAAGTACTTGTCTCGAACTCCCGCGGGGTGTGGGGGTCTCGGCCTCGCT 30
```

```
Db 261 eProSerLeuIleLeuTrpGlyProGlyCysGlyLysThrThrLeuAla 278
RESULT 90
; Sequence 393, Application US/10505928
; Publication No. US20060089532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 393
; LENGTH: 667
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-393

Alignment Scores:
Pred. No.:      8.14      Length:      667
Score:          82.50     Matches:      60
Percent Similarity: 33.7%   Conservative: 31
Best Local Similarity: 22.2% Mismatches: 104
Query Match:     1.9%     Indels:      75
DB:              6       Gaps:       13

US-10-768-158-1 (1-2419) x US-10-505-928-393 (1-667)
Qy 723 CCACCA-----GCTGTGGCAGTCTCCGTCAGGGCTTCAGCTGGGCTTGTGCAC--- 673
Db 28 ProProValSerSerGlyLysAsnGlyProThrSerLeuAlaSerGly-----HisPhe 45
Qy 672 -----AGGACACCCCGAGAAATCTGCCAGCTGTCTCCACCATCG 634
Db 46 ThrGlySerAsnValGluAspArgSerSerGlySerTrpGlyAsnGlyGlyHisPro 65
Qy 633 TCACCAAGT-----CCCGATGCATGCTTTCATCTTGAGAAAAA 595
Db 66 SerProSerArgAsnTyrGlyAspGlyThrProTyrAspHisMetThrSerArgAspLeu 85
Qy 594 GCAGCTTCGAGTCCATCGGTGCTCCAGAACTCTGCACGCTGTCTCAAAACAGGAGCGGT 535
Db 86 GlySerHisAspAsnLeuSerProPheValAsnSerArgIleGlnSerLysThrGlu 105
Qy 534 AGCCAGCTTATCATTAACCTCCGGCAGAAATCTTGAAGGTGC----- 487
Db 106 ArgGlySerTyrSerSerTyrGlyArgGluSerAsnLeuGlnGlyCysHisGlnGlnSer 125
Qy 486 ---CTCGGTAGTTCATGTCGCGCAGAGCGGTGGAACTGATAATAGACACCAACCATG 430
Db 126 LeuLeuGlyGlyAspMetAspMetGlyAsnProGlyThrLeuSerProThrLysPro--- 144
Qy 429 CTTGGGGTTCGAG-----CCATATAGATGACCTTGGAGT 394
Db 145 -----GlySerGlnTyrTyrGlnTyrSerSerAsnAsnProArgArgArgPro----- 160
Qy 393 CTCCATTGTGGAGGTCAGAGGGCAGAAAGCGGTAGGCGAGGTGGCTCTTGTATAGGCGGG 334
Db 161 LeuHisSerSerAlaMetGluValGlnThrLysLysVal----- 173
Qy 333 GAGAGGTGAGTTCCTTGTATGATGTCCAGGCGCGGCTGTGGGTACTCCAGGACCGGGAGCT 274
Db 174 ---ArgLysValPro-----ProGlyLeuProSerSerValTyrAlaProSerAla 189
Qy 273 GCTCGTCGATGTTTCATCAAGCCGATCTCATCGGGGTTCATCGGGGTTCAGCGCTGGCTC 214
Db 190 -----SerThrAlaAspTyrAsnArgAspSerProGlyTyrProSerSer 204
```

Qy	213	CCACCTCTCTGCAGCAAGCTGTGTGCGGACTTTGGGGTAGGTGACGATCCACACGCTCGCTCG	154
Db	205	LysProAlaThr-SerThr-Phe--Pro-----	212
Qy	153	GCGGCACCGGGAAGTTGGCGATCCTCCATCTTCCGCGGCAGAGGGCGGACGCGCA	94
Db	213	-----SerSerPhePheMetGlnAspGlyHisHisSerSerA	225
Qy	93	CGCCATGGAACTCGAAGTACTTGCTCTCGAACTCCCGCGGGTGCTGGGGGTCTCGCGCT	34
Db	225	sppTrpSerSerSerGlyMet---AsnGlnProGlyTyrAlaGlyMetLeuGlyA	244
Qy	33	CGCTCTCCGCCAATGCGCGCGCGCTCG	8
Db	244	snSer-SerHisIleProGlnSerSer	252

```

RESULT 91
US-10-196-749-97
; Sequence 97, Application US/10196749
; Publication No. US20060094864A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P34301C340
; CURRENT APPLICATION NUMBER: US/10/196,749
; CURRENT FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612

```

US-10-768-158-1 (1-2419) x US-10-196-749-97 (1-1894)

759 GGCCACGGGAGGGCCCTCAGCGTTGCAGCACTGGTCCACCA---GCTGGTGGCAGTGCT 703
1019 GlyAlaAlaAlaGlyAlaAlaCysCysThrGlyAlaThrGlyAlaGlyAlaGlyAla 1038
702 CCGTCAGGGCTTCAGCTGGGCCCTTGTACAGGACACCCCGAGGAATCGGCCAGTGCT 643
1039 AlaAlaGlyAlaAlaThrThrGlyCysCysAlaAlaAlaAlaAlaThrThrAlaAla 1058
642 CCACCATCGTCACCGAGTCCCGATGCTCTCATCTTCAGAAAAGACGCTTCGAGT 583
1059 AlaAlaAlaGlyGlyAlaCysAlaCys-----AlaAlaGlyThr 1072
582 CCATCGCGTGTCCAGAACTCTCGACGTCTCAACACGAGGCGTAGCCAGCTTAT 523
1073 GlyCysGlyAlaAlaThrGlyThrAlaAlaAlaThr----- 1085
522 CATTATAAACCTCGGAGAAATCTTGAAGGTGCTCGGTAGCTCATGTGCGCAGAG 463
1086 -----CysAlaGlyCysThrGlyGlyAlaGly 1094
462 AGCGGTGGAATGATAATAAGACACCCAGCATCTTGGGGTTCGAGCCATATAGATGA 403
1095 AlaAlaGly----- 1097
402 CTTGGAGTCTCCATTGTGGAGGTGAGAGGCGAGAGCGTAGGCGAGTGGCTCTTGA 343
1098 -----GlyAlaGlyAlaAlaGlyThrGlyGlyAlaGlyAlaAla--- 1110
342 TGAGCGGGAGAGGTGAGTCTTGTATGTCTCAGGC-----CGGCTGTGGGTACT 289
1111 --GlyAlaAlaAlaThrCysAlaGlyThrCysAlaGlyCysCysGlyCysAlaGlyThr 1129
288 CAGGACCGGGAGCTGCTGCTGATGTTTCATCAAGCCGATCTCATCGGGGTGAGGCCCT 229
1130 GlyAlaAlaGlyAlaGlyCysThrCysAlaGlyAla----- 1141
228 GGCTCACCAGTAGACCACTCTCTGACCAAGTGGTGC CGGACTTGGGGTAGGTGACGA 169
1141 ----- 1141
168 TCCACAGTCGCTGGGCGGACCGGAGTGGCGATCTCTCCATCTTCCCGCGGCGAGA 109
1142 -----AlaAlaAlaGlyAlaAlaGlyCysAlaAlaGlyAlaCysAlaAla 1156
108 AGGGGGGAGCGGACCGCATGGAAGTACTGCTGCTGCAACTCCCCCGGGGTGC 49
1157 ThrThrAlaAlaAlaAlaCysGlyGlyGlyAlaAlaCysThrCysThrThrAlaGlyCys 1176
48 TGGGGGTCTCGGCTCGCTCTCCGCCATCGCGCGCGCGTCCGC 6
1177 -AlaGlyCysAlaAlaAlaAlaCysAlaAlaAlaAlaAla 1190

RESULT 92
US-11-118-524-2
; Sequence 2, Application US/11118524
; Publication No. US2006008847A1
; GENERAL INFORMATION:
; APPLICANT: GU, WEI
; TITLE OF INVENTION: ARF-BP1 AS MEDIATOR OF P53-DEPENDENT AND INDEPENDENT TUMOR
; TITLE OF INVENTION: SUPPRESSION AND USES THEREOF
; FILE REFERENCE: 19240-497US2
; CURRENT APPLICATION NUMBER: US/11/118,524
; CURRENT FILING DATE: 2005-04-29
; PRIOR APPLICATION NUMBER: 60/610,506
; PRIOR FILING DATE: 2004-09-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2
; LENGTH: 4373
; TYPE: PR
; ORGANISM: homo sapiens

US-11-118-524-2									
Alignment Scores:									
Pred. No.:	16.8	Length:	4373						
Score:	82.50	Matches:	88						
Percent Similarity:	35.8%	Conservative:	34						
Best Local Similarity:	25.8%	Mismatches:	119						
Query Match:	1.9%	Indels:	100						
DB:	7	Gaps:	20						
US-10-768-158-1 (1-2419) x US-11-118-524-2 (1-4373)									
QY	832	TGTTATACACCAAGTCAAACTTCATTCATCGACAGCGTGAAG-----ATG	785						
Db	2721	CysThrAlaSerAsnAspSerThrGluGlnAsnLeuSerAspGlyThrProMet	2740						
QY	784	TCCTTCCACAGCCCAACTCTCCCGGCCCGCCAGCGGCTCAGCG-----737							
Db	2741	ProAspSerTyrProThrThrProSerSerThrAspAlaIleThrSerGluSerLysGlu	2760						
QY	736	-----TTGCAGCAGCTGGT-----							
Db	2761	ThrLeuGlyThrLeuGln-SerSerGlnGlnProThrLeuProThrProProAlaLe	2780						
QY	712	TGGCAGTCTCGTCAGGCTT-----CCAGCTGGGCTTGTACAGACACCCCGAGG	659						
Db	2780	uGlyGluValProGlnGluLeuGlnSerProAlaGlyGluGlyGlySerThrGlnLe	2800						
QY	658	AATCTGGCCAGCTGTCCA-----CCATCGTCACAGGTCCCGATCGATCTCT	611						
Db	2800	uLeuMetProValGluProGluGluLeuGlyProThrArgProSerGlyGluAla----	2818						
QY	610	TCATACTTGAGAAAAAGCAGCTTCAGTCCATGCGTCTCCCGA-----AC	563						
Db	2819	-----GluThrGlnMetGluLeuSerProAlaProThrIleThrSerLeuSe	2835						
QY	562	TCCTGCAGCTGTCAACAGAGCGGTAGCCAGCTTATCATTAACAACTCCGCGAG	503						
Db	2835	rProGluArgAlaGluAspSerAspAlaLeuThrAlaValSerSerGlnLeuGluGlySe	2855						
QY	502	A-----ATTCTTGAAGGTGCCTCGGT-----481							
Db	2855	rProMetAspThrSerSerLeuAlaSerCysThrLeuGluGluAlaValGlyAspThrSe	2875						
QY	480	-AGCTCATGTCCGAGAGAGCGGTGAACATGATAATAAGACACCCAGAGT-----CC	428						
Db	2875	rAlaAlaGlySerSerGluGlnProArgAlaGlySerSerThrProGlyAspAlaProPr	2895						
QY	427	TTGGGTTGCGAGCCATATAGATGACCTTGGAGTCTCCATTGTGGAGTCCAGAGGCAGA	368						
Db	2895	oAlaValAlaGlu-----Va	2900						
QY	367	AAGCGTAGGCGAGGTGCTCTTGATAGGCGGGAGAGGTGAGTTCTTGATGATCTCC	308						
Db	2900	lGlnGlyArgSerAspGlySer-----GlyGluSerAlaGlnPro-----Pr	2914						
QY	307	AGGCCCGCTGTGGGTACTCCAGGACCGGAGCTGCTCGTGGATGTTTCATCAAGCCGATC	248						
Db	2914	oGluAspSer-----SerProAlaSerSerGluSerSerThrArgAs	2930						
QY	247	TCATCGGGTTCAGCCCTCGGTACCAAGTAGACCACCTCTCTGAGCAA-----198							
Db	2930	pSerAlaValAlaIleSerGlyAlaAspSerArg-GlyIleLeuGluGluProLeuProS	2950						
QY	197	-----GCTGGTCCCGACTTGGGTAGGTGACGATC	167						
Db	2950	erThrSerSerGluGluAspProLeuAlaGlyIleSerLeuProGluGlyValAspP	2970						
QY	166	CACAC-----GTGCTGGGCG-----CACCGGAAGTT-----GGCGATCTCC	128						
Db	2970	roSerPheLeuAlaAlaLeuProAspAspIleArgArgGluValLeuGlnAsnGlnLeu-	2989						
QY	127	TCCATCTTCCCGGCAGAGGGCGGACGCCATGGCAACTCGAAGTACTTGCTC	68						

US-10-511-937-2466									
; Sequence 2466, Application US/10511937									
; Publication No. US20060088836A1									
; GENERAL INFORMATION:									
; APPLICANT: EXPRESSION DIAGNOSTICS, INC.									
; APPLICANT: Wohlgemuth, Jay									
; APPLICANT: Fry, Kirk									
; APPLICANT: Woodward, Robert									
; APPLICANT: Ly, Ngoc									
; APPLICANT: Prentice, James									
; APPLICANT: Morris, Macdonald									
; APPLICANT: Rosenberg, Steven									
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING									
; TITLE OF INVENTION: AND MONITORING TRANSPLANT REJECTION									
; FILE REFERENCE: 506612000104									
; CURRENT APPLICATION NUMBER: US/10/511,937									
; CURRENT FILING DATE: 2004-10-19									
; PRIOR APPLICATION NUMBER: PCT/US2003/012946									
; PRIOR FILING DATE: 2003-04-24									
; PRIOR APPLICATION NUMBER: US 10/131,831									
; PRIOR FILING DATE: 2002-04-24									
; PRIOR APPLICATION NUMBER: US 10/325,899									
; PRIOR FILING DATE: 2002-12-20									
; NUMBER OF SEQ ID NOS: 3117									
; SOFTWARE: PatentIn version 3.2									
; SEQ ID NO 2466									
; LENGTH: 362									
; TYPE: PRT									
; ORGANISM: Homo sapiens									
US-10-511-937-2466									
Alignment Scores:									
Pred. No.:	7.06	Length:	362						
Score:	82.00	Matches:	65						
Percent Similarity:	26.4%	Conservative:	13						
Best Local Similarity:	22.0%	Mismatches:	82						
Query Match:	1.9%	Indels:	136						
DB:	6	Gaps:	14						
US-10-768-158-1 (1-2419) x US-10-511-937-2466 (1-362)									
QY	29	GAGCCGAGCCGAGACCCC-----CAG	49						
Db	66	GluArgGlyAlaAspAlaThrTrpAspLeuAspLeuLeuLeuThrAsnPheSerGlyPro	85						
QY	50	CACCCCGGGGAGTTCGAGACCAAGTACTTCGAGTTCATGGCTGCGGCTGCCGCTT	109						
Db	86	GluProGlyGlyAlaProGln-----ThrCysAlaLeu	96						
QY	110	CTGCCCGGGAGATGAGGAGATCGCAACTCCCGTGGCGCCAGCGACGTGTGGAT	169						
Db	97	AlaProSerGluAlaSerGlyAlaGlnTyrProPro-----109							
QY	170	CGTCACCTACCCCAAGTCCCGCACCACTTGTCTGCAGGA-----GGTGGTCTA	217						
Db	110	-----ProGluThrLeuGlyAlaGlyProGlyLeuValAlaGlyLeu	126						
QY	218	CTTGGTGAAGCGGCGCTGACCCGATGAGATCGGTTGATGAACATCAACAGTCTCCCG	277						
Db	127	LeuGly-----SerGluAspHisSerGlyTrp	135						
QY	278	CCCGGTCTCGAGTACCCACAGCCGCGCTGGACATCATCAAGAACTGACCTCTCCCG	337						
Db	136	ValArgProAlaLeuArgAlaArgAlaProAspAlaPheValGlyProAlaLeuAlaPro	155						


```
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 872
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-11-302-678-17

Alignment Scores:
Pred. No.: 9.95 Length: 872
Score: 82.00 Matches: 44
Percent Similarity: 32.0% Conservative: 19
Best Local Similarity: 22.3% Mismatches: 60
Query Match: 1.9% Indels: 75
DB: 7 Gaps: 9

US-10-768-158-1 (1-2419) x US-11-302-678-17 (1-872)
QY 1927 ACTCCTCAGATACACGAGGAGTGGAGACACACGACGAGCGGGCGTCCCACTG 1868
Db 672 SerProGluAepSer---ArgGluHisValAspArgHis--GlyCysAileVallysile 689
QY 1867 ATCCGAGT-----GAAACAGGGCTACACTCGCAAAATGTCCTCCACAGGCTCAGCGC 1815
Db 690 ValArgSerSerSerSerThrGlyGlnHisAsnHisPheSerAlaProProAlaAlaProPro 709
QY 1814 CTCAGGTGCACAGAAACAACTGAATGATAAAGACCCGGGGATTTTAGAAATTTACTATTA 1755
Db 710 ValGlnCysProProSerThrSer----- 717
QY 1754 AGCATCTGCTGCCAGCATTAGAGGTATCACTTCAGGAAACAAAGTAAGGCTGTGAA 1695
Db 717 ----- 717
QY 1694 TTCAGCTCAAAATTCGAGCGCAACTCGGTCTTTGGTGGCGGACTCTCCACCCACTG--- 1638
Db 718 -----TrpGlnProGlnSerHisProArgGln 726
QY 1637 ---CATGCAAAATGCTCCT-----GACTGCATTAGACTCCCACTCCAGC 1596
Db 727 GlyHisGlyThrSerProValGlyAspHisGlySerLeuValArgIleProProProPro 746
QY 1595 CTCACGAGGACTTCCCTCAGACATGCTCCGGGGGCTCTCTGCAATG----- 1545
Db 747 ---AlaHisGluArgSerLeuSerAlaTyGlyGlyAsnArgAlaSerMetGluPhe 765
QY 1544 -----GGCCTTTGGAGCCCGCTGGAAGTCTCTGGCAGAAGG 1506
Db 766 LeuArgGlnGluAspThrProGly--CysArgProProGluGlyAsnLeuArgAspSerA 785
QY 1505 AC-----GTCCATTCTGAAGACTAGTGAAACAATTTCTGCTCAT 1467
Db 785 spThrSerIleSerIleProSerValAspHisGluGluGluArgSerPheSerGlyP 805
QY 1466 TTCAGGAAGTTCTGAAGAAGTGAAGTCCGTCAGATAATAACAAGTGT 1420
Db 805 heSerIleSerGlnSerIlyGluAsnLeuAspAlaLeuAsnSerCys 820

RESULT 96
US-10-505-928-54
; Sequence 54, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; PRIOR FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: Patentin 3.2
; SEQ ID NO 54

; LENGTH: 384
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Human Sox-18
US-10-505-928-54

Alignment Scores:
Pred. No.: 7.95 Length: 384
Score: 81.50 Matches: 51
Percent Similarity: 40.1% Conservative: 8
Best Local Similarity: 34.7% Mismatches: 43
Query Match: 1.9% Indels: 46
DB: 6 Gaps: 11

US-10-768-158-1 (1-2419) x US-10-505-928-54 (1-384)
QY 31 GCGAGGCGGACACCC-----CCACACCCCGGGGAGTTCGAGAGCAAGT 75
Db 178 AlaProGlnProProProGluProPheProAlaAlaSerGlySerAlaArgAlaPhe 197
QY 76 ACTTC-----GAGTTCCATGCGGTGCGGCTCCCGCCCTTCTGCCGC 116
Db 198 ArgGluLeuProProLeuGlyAlaGluPheAspGlyLeuGlyLeuProThrProGluArg 217
QY 117 GGAAGATGGAG-----GAGATCGCCAACTTC-----CCGGTGGGCGCC 155
Db 218 SerProLeuAspGlyLeuGluProGlyGlyAlaAlaPhePheProProAlaAlaPro 237
QY 156 AGCCAGCTGTGGATCGTCACCTACCCCAAGTCCG---GCACCAGCTTGTGCAGGA--- 208
Db 238 GluAsp-CysAlaLeuArgProPheArgAlaProTyArgAlaProThr-GluLeuSerArgA 257
QY 209 -----GGTGGTCTACTGGTGAGCCAGGCGCTGACCCCGATGAGATCGGCTTGATGAAC 263
Db 257 spProGlyGlyCysTyArgAlaPro-----LeuAlaGluA 269
QY 264 ATCAGGAGCAGCTCCCGTCTCGAGTACCCACAGCGGG----- 304
Db 269 laLeuArgThrAlaPro---ProAlaAlaProLeuAlaGlyLeuTyArgGlyThrLeuG 288
QY 305 -----CTGGACATCATCAAGGAAGTACCTCTCCCGCTCATCAAGAGCCACT----- 355
Db 288 lyThrProGlyProTyArgProGly---ProLeuSerProProGluAlaProProLeuG 307
QY 356 -----GCCCTACCGCTT 367
Db 307 luSerAlaGluProleu 312

RESULT 97
US-10-511-937-2959
; Sequence 2959, Application US/10511937
; Publication No. US20060088836A1
; GENERAL INFORMATION:
; APPLICANT: EXPRESSION DIAGNOSTICS, INC.
; APPLICANT: Wohlgemuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; APPLICANT: Prentice, James
; APPLICANT: Morris, MacDonald
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; FILE REFERENCE: 50661200104
; CURRENT APPLICATION NUMBER: US/10/511,937
; CURRENT FILING DATE: 2004-10-19
; PRIOR APPLICATION NUMBER: PCT/US2003/012946
; PRIOR FILING DATE: 2003-04-24
; PRIOR APPLICATION NUMBER: US 10/131,831
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US 10/325,899
; PRIOR FILING DATE: 2002-12-20
```

```
; NUMBER OF SEQ ID NOS: 3117
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2959
; LENGTH: 620
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-511-937-2959

Alignment Scores:
Pred. No.:          9.59      Length:      620
Score:              81.50     Matches:    27
Percent Similarity: 44.2%    Conservative: 15
Best Local Similarity: 28.4% Mismatches:   42
Query Match:        1.9%     Indels:     11
DB:                  6       Gaps:        5

US-10-768-158-1 (1-2419) x US-10-511-937-2959 (1-620)
Qy 724 ACCAGTGTGCAACGCTGAGCGCTGCGCGGGCGGGAAGAGTTG---GGCTGTGGA 780
Db 514 SerSerThrGlyThrLysPheProValLysTrpAlaSerProGluValPheSerPheSer 533
Qy 781 AGGACATCTTCACCGCTCTCCATGAATGAGAAAGTTTGACTTGGTGTATAAAC---AGAAGA 837
Db 534 ArgTyrSerSer-----LysSerAspValTrpSerPheGlyValLeuMet 548
Qy 838 TGGGAAAGTGTGACCTCAGCTTGTGACTTTTATATATATACAGAAACAAACCTGCAT 897
Db 549 TrpGlu---ValPheSerGluGlyLysIleProTyrGluAsnArgSerAsnSerGluVal 567
Qy 898 GCTCAATACCCAGACAC-----GTCTACTAGCCAAAGTCTCTGTATGCATTTCATT 948
Db 568 ValGluAspIleSerThrGlyPheArgLeuTyrLysProArgLeuAlaSerThrHisVal 587
Qy 949 TATTCCTTGTGGCAAACTCTGGAAGCAGCGGTGTGAAACACGCG 993
Db 588 TyrGlnIleMetAsnHisCysTrpLysGluArgProGluAspArg 602

RESULT 98
US-10-196-749-466
; Sequence 466, Application US/10196749
; Publication No. US20060094864A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430RIC340
; CURRENT APPLICATION NUMBER: US/10/196,749
; CURRENT FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
```

```
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 466
; LENGTH: 813
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-196-749-466
```

```
Alignment Scores:
Pred. No.:          10.7     Length:      813
Score:              81.50     Matches:    67
Percent Similarity: 35.2%    Conservative: 32
Best Local Similarity: 23.8% Mismatches:   77
Query Match:        1.9%     Indels:     106
DB:                  6       Gaps:        18
```

```
US-10-768-158-1 (1-2419) x US-10-196-749-466 (1-813)
Qy 757 CCCACGGGCGAGGGGCTTCAGCGTTG-----CAGCACTGGTCCACCAGCTGGTGGCAGTGC 704
Db 99  ProbaspGlyGlnProValValLeuAlaProAsnHisThrAspHisCysHisTyrGlnGly 118
Qy 703 TCCGTCTCAGGGCT-----TCCAGCTGGGCTTGTACAGGACACACCCCGGAAATCTGGCC 650
Db 119 ArgValArgGlyPheProAspSerTrpValVal-----LeuCys 131
Qy 649 AGCTGC-----TCCACCATCTGCACAGGTCCCGATGCATGCTTTCATATTGAGA 599
Db 132 ThrCysSerGlyMetSerGlyLeuIleThrLeuSerArgAsnAlaSerTyrTyrLeuArg 151
Qy 598 AAA-----AGCACGTTCCGATCCATCGCGTGTCTCCAG 566
Db 152 ProTrpProProArgGlySerLysAspPheSerThrHisGluIlePheArgMetGluGln 171
Qy 565 AAC-----TCCTGCACGTGCTCAACACGAGCGGTAGCCGACGCTTATCATTCATA 515
Db 172 LeuLeuThrTrpLysGlyThrCysGlyHisArgAspProGlyAsnLysAlaGlyMetThr 191
Qy 514 AACCTCCGGCAGAAATCTTG-----AAAGTCCCTCGGTAGCTCATGTCGCCAGAGCG 459
Db 192 SerLeuProGlyGlyProGlnSerArgGlyArgGluAlaArgArgThrArgLysTyr 211
Qy 458 GTGGAACCTGATAATA-----AGACAC----- 438
Db 212 LeuGluLeuTyrIleValAlaAspHisThrLeuPheLeuThrArgHisArgAsnLeuAsn 231
Qy 437 -----CACAGATCCTTGGGTTGCGAGCCATATAGATGACCTGGAGTCTCCATT 387
Db 232 HisThrLysGlnArgLeuLeuGluValAlaAsnTyrValAspGlnLeu---LeuArgThr 250
Qy 386 GTGGAGGTACAGGGCAGAAAGCGTAGGCGAGTGGCTCTT----- 345
Db 251 LeuAspIle-----GlnValAlaLeuThrGlyLeuGluValTrp 263
Qy 344 GATCAGCGGGGAGAGGTGCTTCTTGTATGATGTCAGAGCGCGCTGTGGTA---CTC 288
Db 264 ThrGluArgAspArgSerArgValThrGlnAspAlaAsnAlaThrLeuTrpAlaPheLeu 283
Qy 287 CAGGACCGG-----GAGCTCTCTC----- 270
Db 284 GlnTrpArgGlyLeuTrpAlaGlnArgProHisAspSerAlaGlnLeuLeuThrGly 303
Qy 270 ----- 270
Db 304 ArgAlaPheGlnGlyAlaThrValGlyLeuAlaProValGluGlyMetCysArgAlaGlu 323
Qy 269 -----GTCGATGTTTCATCAAGCCGATCTC---ATCGGGGTGAGCGCC---CTG 228
```



```
Db      1103 erGlnPheLeuGlySerAlaThrSerProGlyProArgThrAlaProAlaAlaLysProG 1123
      |||::: |||
Qy      867 AAAAGTCAACGTGAGGTCAACATTTCCATCTCTCTTTATACACCAAGTCAAACTTC- 809
      |||::: |||
Db      1123 lulyaSaSerAlaGlyThrArgValProVal-----GlnGluLysThrAsnLeuA 1140
      |||::: |||
Qy      808 --TCATTATGAGACGGTGAAGATGTCCTTCCACAGCCCAACTCTTCCCGGCCACGG 751
      |||::: |||
Db      1140 laAlaTyValProLeuLeuThrGlnGlyTrpAlaGluLeuValArgArgProThrG 1160
      |||::: |||
Qy      750 GCAGGGCTCAGCGTTG----- 734
      |||
Db      1160 lyAsnThrSerTrpLeuMetSerLeuGluAsnProLeuSerProPheSerSerAspIleA 1180
      |||
Qy      733 -----CAGC 730
      |||
Db      1180 snAsnMetProLeuGlnGluLeuSerAsnAlaLeuMetAlaAlaGluArgPheLysGluH 1200
      |||
Qy      729 ACTGGTCCACAGCTGGTGGCAG--TGCTCCGTCAGGGCTTCCAGCTGGGCTTGTTCAC 673
      |||::: |||
Db      1200 isArgAspThrAlaLeuTyrlsSerLeuSerValProAlaAlaSerThrAlaLysProp 1220
      |||::: |||
Qy      672 AGGACACCCCGAGAAAT-----CTGGCCAGCTGCTCCACCATCGTCACCGGCCCGAT 619
      |||::: |||
Db      1220 roProLeuProArgSerAsnThrValAlaSerPheSerSerLeuTyR----GlnSerSerC 1239
      |||::: |||
Qy      618 GCATGTCTTCATACTTGAGAAAAGCAGTTC----- 587
      |||
Db      1239 ysGlnGlyGlnLeuHisArgSerValSerTrpAlaAspSerAlaValMetGluGluG 1259
      |||
Qy      586 -----GAGTCCATCGGTGCTCCAGAACTCCTGCA 556
      |||
Db      1259 lySerProGlyGluValProValLeuValGlu-ProproGlyLeuGluAspValGluAla 1278
      |||
Qy      555 CGTGCTCAACACGAGCGGTAGCCAGCTTATCATTCATAAAGCTCC----- 508
      |||
Db      1279 AlaLeuGlyMetAspArgArgThrAspAlaTyrlsSerArgSerSerValSerSerGln 1298
      |||
Qy      507 -----GCAGAAATTTGAAGGTGC 487
      |||
Db      1299 GluGluLysSerLeuHisAlaGluLeuValGlyArgGlyIleProIleGluArgVal 1318
      |||
Qy      486 CTCGGTAGCTCATGTCGCGAGAGCGGTGGAACTGATATAGACACCAACCATCCT 427
      |||
Db      1319 ValSerSer-GluGlyArgProSerValAspLeuSerPheGlnProSerGlnProLe 1338
      |||
Qy      426 TGGGGTTCCGAGCCATATAGATACCTTGGAGTCTCCATTGTGGAGGTCCAGGGCGAGAA 367
      |||
Db      1338 uSerLysSerSerSerProGluLeu--GlnThrLeuGlnAspIleLeuGlyAspPr 1357
      |||
Qy      366 AGCGGTAGGCGAGGTGCTTGTATGAGCGGGGAGAGGTGATGCTTGTATGATGTCCA 307
      |||
Db      1357 ogly-----AspLysAlaAspValGlyArgLeuSerProGluVally 1371
      |||
Qy      306 GGCCCGGCT---GTGGTACTCCAGGACCGGGAGCTGCTCGATCTTTCATCAAGCGC 251
      |||
Db      1371 sAlaArgSerGlnSerGlyThrLeuAspGlyGluSerAlaAlaTrpSerAlaSerGlyG 1391
      |||
Qy      250 ATCTCATCGGGGTGAGCCCGCTGGTCTCA---CCAAGTAGACCACTCTCTGCAAGAGCTG 194
      |||
Db      1391 uAspSerArgGlyGlnProGluGlyProLeuProSerSerSerProArgSerProSerG 1411
      |||
Qy      193 GTGCCGAGTGGGGTAGTGACGATCCACAGCTGCTGGGCGGCGACCGCCACCGATGGAAGTGGCG 134
      |||
Db      1411 yLeuArgProArgGlyTyrlsSerAsp----- 1421
      |||
Qy      133 ATCTCCTCCATCTTCCCGCGCAGAGGGCGGCGACCGCCACCGATGGAAGTCCGAGTAC 74
      |||
Db      1422 -SerAlaProSerArgArgGlyLysArgValGluArgAspAlaLeuLysSerArgAlaTh 1441
      |||
Qy      73 TTGCTCTCGAACTCCC-----CGGGG 52
      |||
```

```
Db      1441 rAlaSerAsnAlaGluLysValProGly 1450
      |||
RESULT 100
US-10-196-749-224
; Sequence 224, Application US/10196749
; Publication No. US20060094864A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C340
; CURRENT APPLICATION NUMBER: US/10/196,749
; PRIOR FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 224
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-196-749-224
Alignment Scores:
Pred. No.: 8.36 Length: 341
Score: 81.00 Matches: 51
Percent Similarity: 33.7% Conservative: 17
Best Local Similarity: 25.2% Mismatches: 61
Query Match: 1.8% Indels: 74
DB: 6 Gaps: 13
US-10-768-158-1 (1-2419) x US-10-196-749-224 (1-341)
Qy      594 GCAGTTTCGAGTCCATCGGTGCTCCAGAACTCCTCCACGTGCTCAACCCAGG----- 541
      |||
Db      121 SerArgGluLeuProSerAlaThrProAsnThrAlaGlySerSerThrArgPheIle 140
      |||
Qy      540 -----AGCGGTAGCCAGCTTATCATTAACCTCC-----GGCAGA 502
      |||
Db      141 AlaAsnSerGlnGluProGluIleArgLeuThrSerSerLeuProAspSerProGlyArg 160
      |||
Qy      501 ATTCTTGAAGGTGCCTCGGT-----AGCTCATGCTCGCAGAGCGGT 457
      |||
Db      161 SerThrGluAspLeuProGlySerGlnAlaThrLeuSerGlnTrpSerThr---ProGly 179
      |||
```

```
Qy 456 GGAAGTATATAGACACCA-----CCAGAT----- 430
Db 180 SerThrProSerArgTrpProSerProSerProThrAlaMetProSerProGluAspLeu 199
Qy 429 -----CCTTGGGTTGCCAGCATATAGATGACCTTGGAGTCTCCATTGT 385
Db 200 ArgLeuValLeuMetProTrpGly-----ProTrpHisCysHisCys 213
Qy 384 GGAGGT-----CAGAGGCGAGAAAGCGGTAGGCGAGGTGGCTC-----TTGATG 341
Db 214 LysSerGlyThrMetSerArgSerArgSerGlyLysLeuHisGlyLeuSerGlyArgLeu 233
Qy 340 AGGCGGGGAGAGGTCCCTTGTATGATGATGCCAGGCCGGCTGTGGTACTCCAGG---- 284
Db 234 ArgValGlyAlaLeuSerGlnLeuArgThrGluHisLysProCysThrTyrGlnGlnCys 253
Qy 283 -----ACCGGGAGC 275
Db 254 ProCysAsnArgLeuArgGluGluCysProLeuAspThrSerLeuCysThrAspThrAsn 273
Qy 274 TGCTCGTCGATGTTTCATCAAGCCGATCTCATCGGGTCAGCGCCCT----- 229
Db 274 CysAlaSerGln-SerThrThrSerThrArgThrThrThrThrProPheProThrIleHi 293
Qy 228 -----GGCTCACCAGTAGACCCT-----CCTGCA-----GCAAGCTGGTG 191
Db 293 sLeuArgSerSerProSerLeuProProAlaSerProCysProAlaLeuAlaPheTrpLy 313
Qy 190 CCGG 187
Db 313 sArg 314
```

Search completed: May 17, 2006, 10:59:18
Job time : 313 secs

THIS PAGE BLANK (USPTO)

85 289.5 6.6 312 4 US-10-072-012-665 Sequence 665, App
 86 281.5 6.4 303 4 US-10-072-012-250 Sequence 250, App
 87 281 6.4 317 4 US-10-437-963-157482 Sequence 157482,
 88 277 6.3 345 6 US-11-096-568A-1016 Sequence 1016, Ap
 89 275.5 6.3 320 3 US-09-854-122-20 Sequence 20, App1
 90 275.5 6.3 320 3 US-11-033-030-20 Sequence 20, App1
 91 275 6.2 340 6 US-11-096-568A-12566 Sequence 12566, A
 92 274 6.2 282 6 US-11-096-568A-1018 Sequence 1018, Ap
 93 274 6.2 286 6 US-11-096-568A-1017 Sequence 1017, Ap
 94 269.5 6.1 336 4 US-10-437-963-198983 Sequence 198983,
 95 268.5 6.1 329 6 US-11-096-568A-7712 Sequence 7712, Ap
 96 268.5 6.1 330 6 US-11-096-568A-7711 Sequence 7711, Ap
 97 268.5 6.1 340 6 US-11-096-568A-7710 Sequence 7710, Ap
 98 267.5 6.1 344 4 US-10-767-701-42402 Sequence 42402, A
 99 267 6.1 343 4 US-10-259-194A-330 Sequence 330, App
 100 267 6.1 343 4 US-10-437-963-159374 Sequence 159374,

ALIGNMENTS

RESULT 1

US-10-427-631-7
 ; Sequence 7, Application US/10427631
 ; Publication No. US20030175923A1
 ; GENERAL INFORMATION:
 ; APPLICANT: INCYTE CORPORATION; TANG, Y. Tom;
 ; APPLICANT: CORLEY, Neil C.; GUEGLER, Karl J.;
 ; APPLICANT: BAUGHN, Mariah R.; LAL, Preeti G.;
 ; APPLICANT: YUE, Henry; HILLMAN, Jennifer L.;
 ; APPLICANT: AZIMZAI, Yalda
 ; TITLE OF INVENTION: HUMAN TRANSFERASE PROTEINS
 ; FILE REFERENCE: PF-0592-1 DIV
 ; CURRENT APPLICATION NUMBER: US/10/427,631
 ; CURRENT FILING DATE: 2003-04-29
 ; PRIOR APPLICATION NUMBER: US 09/786,240
 ; PRIOR FILING DATE: 2002-03-12
 ; PRIOR APPLICATION NUMBER: PCT/US99/20989
 ; PRIOR FILING DATE: 1999-09-09
 ; PRIOR APPLICATION NUMBER: US 60/172,220
 ; PRIOR FILING DATE: 1998-09-10
 ; PRIOR APPLICATION NUMBER: US 60/155,248
 ; PRIOR FILING DATE: 1998-11-04
 ; PRIOR APPLICATION NUMBER: US 60/133,642
 ; PRIOR FILING DATE: 1999-05-11
 ; NUMBER OF SEQ ID NOS: 33
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 7
 ; LENGTH: 284
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: Incyte ID No. US20030175923A1 1784742CD1
 US-10-427-631-7

Alignment Scores:
 Pred. No.: 1.71e-119 Length: 284
 Score: 1530.00 Matches: 284
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 34.8% Indels: 0
 Db: 4 Gaps: 0

US-10-768-158-1 (1-2419) x US-10-427-631-7 (1-284)

QY 21 ATGGGGAGAGCGAGCGCGGAGTTCGAGAGCAAGTACTTC 80
 |||||
 Db 1 MetAlaGluSerGluAlaGluThrProSerThrProGlyGluPheGluSerIysTyrPhe 20
 |||||
 QY 81 GAGTTCCATGGCGTCCGCGCTCTCTGCGCGGGAAGATGGAGGATCGCCAAAC 140
 |||||
 Db 21 GluPheHisGlyValArgLeuProPhePheCysArgGlyLysMetGluGluIleAlaAsn 40
 |||||

QY 141 TTCCCGGTGGCGCCAGCGACGTGTGGATCGTACCTACCCCAAGTCCGGCACCAGCTTG 200
 |||||
 Db 41 PheProValArgProSerAspValTrpIleValThrTyrProLysSerGlyThrSerLeu 60
 |||||
 QY 201 CTGCAGGAGGTGGTCTACTTGGTGAGCGAGGCGCTGACCCCGATGAGATCGGCTTGATG 260
 |||||
 Db 61 LeuGlnGluValValTyrLeuValSerGlnGlyAlaAspProAspGluIleGlyLeuMet 80
 |||||
 QY 261 AACATCAGCAGCAGCTCCCGGTCTGAGTACCCAGCGCGGCGCTGGACATCATCAAG 320
 |||||
 Db 81 AsnIleAspGluGlnLeuProValLeuGluTyrProGlnProGlyLeuAspIleIleLys 100
 |||||
 QY 321 GAATGACCTCTCCCGCGCTCATCAAGAGCACCCTGCTACCGCTTTCTGCGCTCTGAC 380
 |||||
 Db 101 GluLeuThrSerProArgLeuIleLysSerHisLeuProTyrArgPheLeuProSerAsp 120
 |||||
 QY 381 CTCCACAATCGAGACTCCAAGGTCTATATGGCTCGCAACCCCAAGGATCTCGTGGTG 440
 |||||
 Db 121 LeuHisAsnGlyAspSerLysValIleTyrMetAlaArgAsnProLysAspLeuValVal 140
 |||||
 QY 441 TCTTATTATCAGTTCCACCGCTCTCTCGGACCATGAGTACCGAGGCACCTTTCAGAA 500
 |||||
 Db 141 SerTyrTyrGlnPheHisArgSerLeuArgThrMetSerTyrArgGlyThrPheGlnGlu 160
 |||||
 QY 501 TTCTCGCGGAGGTTTATGAATGATAAGCTGGGCTACGCTCTCTGGTTTGGACGCTGCAG 560
 |||||
 Db 161 PheCysArgArgPheMetAsnAspLysLeuGlyTyrGlySerTrpPheGluHisValGln 180
 |||||
 QY 561 GAGTTCTGGGAGCACCCTGAGTTCGAACGTGCTTTTCTCAAGTATGAAGACATGCAT 620
 |||||
 Db 181 GluPheTrpGluHisArgMetAspSerAsnValLeuPheLeuLysTyrGluAspMetHis 200
 |||||
 QY 621 CGGACCTGTGAGCATGTTGGAGCAGCTGCCAGATTCCTGGGGGTCTCTGTGACAAAG 680
 |||||
 Db 201 ArgAspLeuValThrMetValGluGlnLeuAlaArgPheLeuGlyValSerCysAspLys 220
 |||||
 QY 681 GCCCAGCTGGAAGCCCTGACGAGCAGCTGCCACAGCTGGTGGACAGTGTCTGCAACGCT 740
 |||||
 Db 221 AlaGlnLeuGluAlaLeuThrGluHisCysHisGlnLeuValAspGlnCysAsnAla 240
 |||||
 QY 741 GAGCCCTGCGCGTGGCGCGGGAAGAGTTGGGCTGTGGAGGACATCTTCCACGCTCTCC 800
 |||||
 Db 241 GluAlaLeuProValGlyArgGlyArgValGlyLeuTrpLysAspIlePheThrValSer 260
 |||||
 QY 801 ATGAATCAGAGTTTGGCTGTATATAACAGAGATGGAAAGTGTGACCTCAGCTTT 860
 |||||
 Db 261 MetAsnGluLysPheAspLeuValTyrLysGlnLysMetGlyLysCysAspLeuThrPhe 280
 |||||
 QY 861 GACTTTTATTATA 872
 |||||
 Db 281 AspPheTyrLeu 284

RESULT 2

US-10-757-262-110
 ; Sequence 110, Application US/10757262
 ; Publication No. US20040197825A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Karicheti, Venkateswarlu
 ; APPLICANT: Silos-Santiago, Inmaculada
 ; APPLICANT: Eliasof, Scott D.
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
 ; TITLE OF INVENTION: UROLOGICAL DISORDERS USING 44390, 54181, 211, 5687, 884,
 ; TITLE OF INVENTION: 1405, 636, 4421, 5410, 30905, 2045, 15405, 18560, 2047,
 ; TITLE OF INVENTION: 33751, 52872, 14063, 20739, 32544, 43239, 44373, 51164,
 ; TITLE OF INVENTION: 53010, 16852, 1587, 2207, 22245, 2387, 52908, 69112, 14990,
 ; TITLE OF INVENTION: 18047, 115, 579, 15985, 15625, 760, 18603, 2395, 2554, 8675,
 ; TITLE OF INVENTION: 32720, 4809, 14303, 16816, 17827, 32620, 577, 619, 1423,
 ; TITLE OF INVENTION: 2158, 8263, 15402, 16209, 16386, 21165, 30911, 41897, 1643,
 ; TITLE OF INVENTION: 2543, 9626, 13231, 32409, 84260, 2882, 8203, 32678 OR
 ; TITLE OF INVENTION: 55053
 ; FILE REFERENCE: MP103-007P1RNMNIM
 ; CURRENT APPLICATION NUMBER: US/10/757,262
 ; CURRENT FILING DATE: 2004-01-14

;; PRIOR APPLICATION NUMBER: US 60/440,318
;; PRIOR FILING DATE: 2003-01-15
;; PRIOR APPLICATION NUMBER: US 60/444,783
;; PRIOR FILING DATE: 2003-02-04
;; PRIOR APPLICATION NUMBER: US 60/457,901
;; PRIOR FILING DATE: 2003-03-27
;; PRIOR APPLICATION NUMBER: US 60/468,775
;; PRIOR FILING DATE: 2003-05-08
;; PRIOR APPLICATION NUMBER: US 60/471,614
;; PRIOR FILING DATE: 2003-05-19
;; PRIOR APPLICATION NUMBER: US 60/478,742
;; PRIOR FILING DATE: 2003-06-16
;; PRIOR APPLICATION NUMBER: US 60/488,529
;; PRIOR FILING DATE: 2003-07-18
;; PRIOR APPLICATION NUMBER: US 60/491,156
;; PRIOR FILING DATE: 2003-07-30
;; PRIOR APPLICATION NUMBER: US 60/499,594
;; PRIOR FILING DATE: 2003-09-02
;; PRIOR APPLICATION NUMBER: US 60/506,332
;; PRIOR FILING DATE: 2003-09-26
;; NUMBER OF SEQ ID NOS: 136
;; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 110
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-757-262-110

Alignment Scores:
Pred. No.: 1.71e-119 Length: 284
Score: 1530.00 Matches: 284
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 34.8% Indels: 0
DB: 4 Gaps: 0

US-10-768-158-1 (1-2419) x US-10-757-262-110 (1-284)

Qy 21 ATGGCGGAGAGCGAGCGGAGACCCCGGAGACCCCGGCGGAGTTCGAGAGCAAGTACTTC 80
Db 1 MetAlaGluSerGluAlaGluThrProSerThrProGlyGluPheGluSerIleTyPhe 20
Qy 81 GAGTTCCATGCGTGGCGTCCGCGCTTCGCGCGGGAAGATGGAGAGATCGCCCAAC 140
Db 21 GluPheHisGlyValArgLeuProPheCysArgGlyLysMetGluIleAlaAsn 40
Qy 141 TTCGGTGGCGGCGGCGGAGCGAGTGGATGTCACCTACCCCAAGTCCGGACCGCTTG 200
Db 41 PheProValArgProSerAspValTrpIleValThrTyProLysSerGlyThrSerLeu 60
Qy 201 CTGCAGGAGGTGGTCTACTTGTGAGCCAGGCGGTGACCCCGATGAGATCGGCTTGATG 260
Db 61 LeuGlnGluValValTyLeuValSerGlnGlyAlaAspProAspGluIleGlyLeuMet 80
Qy 261 AACATCGACGAGCTCCCGTCTCTGGAGTACCCACAGCGCGGCTCGACATCATCAAG 320
Db 81 AsnIleAspGluGlnLeuProValLeuGluTyProGlnProGlyLeuAspIleIleLys 100
Qy 321 GAATGACCTCTCCCGCTCATCAAGAGCCACCTGCGCTACCGCTTCTGCGCTCTGAC 380
Db 101 GluLeuThrSerProArgLeuIleLysSerHisLeuProTyArgPheLeuProSerAsp 120
Qy 381 CTCACAATGAGAGCTCCAAGTCTATCATATGGTTCGCAACCCCAAGGATCTGGTGTG 440
Db 121 LeuHisAsnGlyAspSerLysValIleTyMetAlaArgAsnProLysAspLeuVal 140
Qy 441 TCTTATTATCAGTTCACCGCTCTCTGGGAGCCATGAGCTACCGAGGACCTTTCAAGAA 500
Db 141 SerTyTyGlnPheHisArgSerLeuArgThrMetSerTyArgGlyThrPheGlnGlu 160
Qy 501 TTTCCCGGAGTTTATGAATGATTAAGCTGGGCTACGGCTCTGTTTGGACAGCTGAG 560
Db 161 PheCysArgPheMetAsnAspLysLeuGlyTyGlySerTrpPheGluHisValGln 180

Qy 561 GAGTTCTGGGAGACCCGATGGACTCGAAGCGTCTTTTCTCAAGTATGAAGACATGCAT 620
Db 181 GluPheTrpGluHisArgMetAspSerAsnValLeuPheLeuLysTyGluAspMetHis 200
Qy 621 CGGACCTGCTGAGATGGTGGACGACGCTGCCAGATTCCTGGGGGTCTCTGTGACAAG 680
Db 201 ArgAspLeuValThrMetValGluGlnLeuAlaArgPheLeuGlyValSerCysAspLys 220
Qy 681 GCCCAGCTGGAAGCCCTGACGGAGCACTGCCACGAGCTGGTGGACCAAGTCTGCAACGCT 740
Db 221 AlaGlnLeuGluAlaLeuThrGluHisCysHisGlnLeuValAspGlnCysCysAsnAla 240
Qy 741 GAGCCCTGCCCGTGGCGGCGGGAAGAGTTGGGCTGTGGAGAGCATCTTCACCGTCTCC 800
Db 241 GluAlaLeuProValGlyArgGlyArgValGlyLeuTrpLysAspIlePheThrValSer 260
Qy 801 ATGATCAGAGAGTTGACTTGGTGTATATAACAGAGATGGAAAGTGTGACCTCACGTTT 860
Db 261 MetAsnGluLysPheAspLeuValTyLysGlnLysMetGlyLysCysAspLeuThrPhe 280
Qy 861 GACTTTTATTATA 872
Db 281 AspPheTyLeu 284

RESULT 3

US-10-768-158-2
; Sequence 2, Application US/10768158
; Publication No. US20040204359A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Silos-Santiago, Inmaculada
; APPLICANT: Karicheti, Venkateswarlu
; APPLICANT: Eliasof, Scott D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: PAIN AND PAINFUL DISORDERS USING 16386, 15402, 21165, 1423,
; TITLE OF INVENTION: 636, 12303, 21425, 27410, 38554, 38555, 55063, 57145, 59914,
; TITLE OF INVENTION: 94921, 16852, 33260, 58573, 30911, 85913, 14303, 16816,
; TITLE OF INVENTION: 17827 OR 32620
; FILE REFERENCE: MPI03-012P1RNONNIM
; CURRENT APPLICATION NUMBER: US/10768,158
; CURRENT FILING DATE: 2004-01-30
; PRIOR APPLICATION NUMBER: US 60/444,781
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: US 60/452,291
; PRIOR FILING DATE: 2003-03-05
; PRIOR APPLICATION NUMBER: US 60/454,540
; PRIOR FILING DATE: 2003-03-13
; PRIOR APPLICATION NUMBER: US 60/478,805
; PRIOR FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: US 60/491,048
; PRIOR FILING DATE: 2003-07-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-768-158-2

Alignment Scores:
Pred. No.: 1.71e-119 Length: 284
Score: 1530.00 Matches: 284
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 34.8% Indels: 0
DB: 4 Gaps: 0

US-10-768-158-1 (1-2419) x US-10-768-158-2 (1-284)

Qy 21 ATGGCGGAGAGCGAGCGGAGACCCCGGAGACCCCGGCGGAGTTCGAGAGCAAGTACTTC 80
Db 1 MetAlaGluSerGluAlaGluThrProSerThrProGlyGluPheGluSerLysTyPhe 20

```
Qy 81 GAGTTCCATGCGGTGCGGCTGCCGCTTCTGCGCGGAGAGATGGAGGATCGCCAAC 140
Db 21 GluPheHisGlyValArgLeuProPheCysArgGlyLysMetGluGluValAlaAsn 40
Qy 141 TTCCCGGTGCGGCGGAGCGAGTGTGATCTGATCTACCTACCCCAAGTCCGGCACCAGCTTG 200
Db 41 PheProValArgProSerAspValTrpIleValThrTyrProLysSerGlyThrSerLeu 60
Qy 201 CTGACGAGGTGTCTACTTGGTGAGCCAGGCGCTGACCCCGATGAGATCGGCTTGATG 260
Db 61 LeuGlnGluValValTyrLeuValSerGlnGlyAlaAspProAspGluIleGlyLeuMet 80
Qy 261 AACATCCAGCAGCAGCTCCCGTCTCTGAGTACCCACAGCGCGCTGGACATCATCAAG 320
Db 81 AsnIleAspGluGlnLeuProValLeuGluTyrProGlnProGlyLeuAspIleIleLys 100
Qy 321 GAATGACCTCTCCCGCTCATCAAGAGCCACCTGCCCTACCGCTTTCTGCCCTCTGAC 380
Db 101 GluLeuThrSerProArgLeuIleLysSerHisLeuProTyrArgPheLeuProSerAsp 120
Qy 381 CTCACAAATGAGACTCAAGGTCAATATATGCTGCAACCCCAAGGATCTGGTGGTG 440
Db 121 LeuHisAsnGlyAspSerLysValIleTyrMetAlaArgAsnProLysAspLeuVal 140
Qy 441 TCTTATTATCAGTTCACCGCTCTCTCGGACCATGAGCTACCGAGGCACCTTTCAAGAA 500
Db 141 SerTyrTyrGlnPheHisArgSerLeuArgThrMetSerTyrArgGlyThrPheGlnGlu 160
Qy 501 TTCTGCCGAGGTTTATGAATGATAAGCTGGGCTACGGCTCCTGGTTTGACACGTCGAG 560
Db 161 PheCysArgArgPheMetAsnAspLysLeuGlyTyrGlySerTrpPheGluHisValGln 180
Qy 561 GAGTTCTGGAGCAGCATCGAATGATGATGATGATGATGATGATGATGATGATGATGAT 620
Db 181 GluPheTrpGluHisArgMetAspSerAsnValLeuPheLeuLysTyrGluAspMetHis 200
Qy 621 CGGACCTGGTGAGCAGTGGTGGAGCAGCTGCCACAGCTGGTGAGCAGCTGGTGGAG 680
Db 201 ArgAspLeuValThrMetValGluGlnLeuAlaArgPheLeuGlyValSerCysAspLys 220
Qy 741 GAGGCTCTGCCGCGGAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 740
Db 241 GluAlaLeuProValGlyArgGlyValGlyLeuTrpLysAspIlePheThrValSer 260
Qy 801 ATGAATGAGAAGTTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 860
Db 261 MetAsnGluLysPheAspLeuValTyrLysGlnLysMetGlyLysCysAspLeuThrPhe 280
```

RESULT 4

```
US-10-205-331-46
; Sequence 46, Application US/10205331
; Publication No. US20040058326A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; APPLICANT: Lee, Kevin
; APPLICANT: Dixon, Alistair
; APPLICANT: Brooksbank, Robert
; APPLICANT: Pinnock, Robert
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
; FILE REFERENCE: WL-A-018199
; CURRENT APPLICATION NUMBER: US/10/205,331
; PRIOR FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: GB 0118354.0
; PRIOR FILING DATE: 2001-07-27
```

```
; NUMBER OF SEQ ID NOS: 117
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 46
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Sulfotransferase-like protein
US-10-205-331-46
```

```
Alignment Scores:
Pred. No.: 1.79e-117 Length: 284
Score: 1506.00 Matches: 278
Percent Similarity: 99.3% Conservative: 4
Best Local Similarity: 97.9% Mismatches: 2
Query Match: 34.2% Indels: 0
DB: 4 Gaps: 0
```

US-10-768-158-1 (1-2419) x US-10-205-331-46 (1-284)

```
Qy 21 ATGCGGAGAGCGGAGCGGAGACCCCGAGCAGCCCGGGGAGTTCGAGAGCAAGTACTTC 80
Db 1 MetAlaGluSerGluAlaGluThrProGlyThrProGlyGluPheGluSerLysfyrPhe 20
Qy 81 GAGTTCCATGCGGTGCGGCTGCCGCTTCTGCGCGGAGAGATGGAGGATCGCCAAC 140
Db 21 GluPheHisGlyValArgLeuProPheCysArgGlyLysMetGluAspIleAlaAsp 40
Qy 141 TTCCCGGTGCGGCGGAGCGAGTGTGATCGTCACTACCTACCCCAAGTCCGGCACCAGCTTG 200
Db 41 PheProValArgProSerAspValTrpIleValThrTyrProLysSerGlyThrSerLeu 60
Qy 201 CTCGAGCAGTGTGTACTTGTGAGCCAGGCGCTGACCCCGATGAGATCGGCTTGATG 260
Db 61 LeuGlnGluValValTyrLeuValSerGlnGlyAlaAspProAspGluIleGlyLeuMet 80
Qy 261 AACATCCAGCAGCAGCTCCCGCTCTCGAGTACCCACAGCGGCGCTGGACATCATCAAG 320
Db 81 AsnIleAspGluGlnLeuProValLeuGluTyrProGlnProGlyLeuAspIleIleLys 100
Qy 321 GAATGACCTCTCCCGCTCATCAAGAGCAGCTGCCCTACCGCTTTCTGCCCTCTGAC 380
Db 101 GluLeuThrSerProArgLeuIleLysSerHisLeuProTyrArgPheLeuProSerAsp 120
Qy 381 CTCACAAATGAGACTCAAGGTCAATATATGCTGCAACCCCAAGGATCTGGTGGTG 440
Db 121 LeuHisAsnGlyAspSerLysValIleTyrMetAlaArgAsnProLysAspLeuVal 140
Qy 441 TCTTATTATCAGTTCACCGCTCTCTCGGACCATGAGCTACCGAGGCACCTTTCAAGAA 500
Db 141 SerTyrTyrGlnPheHisArgSerLeuArgThrMetSerTyrArgGlyThrPheGlnGlu 160
Qy 501 TTCTGCCGAGGTTTATGAATGATGATGATGATGATGATGATGATGATGATGATGAT 560
Db 161 PheCysArgArgPheMetAsnAspLysLeuGlyTyrGlySerTrpPheGluHisValGln 180
Qy 561 GAGTTCTGGAGCAGCATCGAATGATGATGATGATGATGATGATGATGATGATGATGAT 620
Db 181 GluPheTrpGluHisArgMetAspSerAsnValLeuPheLeuLysTyrGluAspMetHis 200
Qy 621 CGGACCTGGTGAGCAGTGGTGGAGCAGCTGCCACAGCTGGTGAGCAGCTGGTGGAG 680
Db 201 ArgAspLeuValThrMetValGluGlnLeuAlaArgPheLeuGlyValSerCysAspLys 220
Qy 681 GCCAGCTGGAGCGCTGAGCGGACCATGCGCAGCTGCCACAGCTGGTGGACCGAGTGTG 740
Db 221 AlaGlnLeuGluSerLeuIleGluHisCysHisGlnLeuValAspGlnCysCysAsnAla 240
Qy 741 GAGGCTCTGCCGCGGAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 800
Db 241 GluAlaLeuProValGlyArgGlyValGlyLeuTrpLysAspIlePheThrValSer 260
Qy 801 ATGAATGAGAAGTTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 860
```



```
Db      261 MetAsnGluLysPheAspLeuValIrrLysGlnLysMetGlyLysCysAspLeuThrPhe 280
Oy      861 GACTTTTATTTA 872
Db      281 AspPheTyrLeu 284

RESULT 5
US-09-989-442-113
; Sequence 113, Application US/09989442
; Publication No. US20030013649A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P0208
; CURRENT APPLICATION NUMBER: US/09/989,442
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/225,757
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/226,868
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/216,647
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,267
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/216,880
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,270
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/251,869
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/235,834
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/234,274
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/234,223
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/228,924
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/224,518
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,369
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/224,519
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,964
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/241,809
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/249,299
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/236,327
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/241,785
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/244,617
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/225,268
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,368
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/251,856
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/251,868
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/229,344
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/234,997
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: 60/229,343
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,345
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,287
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,513
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/231,413
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/229,509
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/236,367
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/237,039
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,038
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/236,370
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/236,802
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,037
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,040
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/240,960
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/239,935
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/239,937
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/241,787
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,474
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/246,532
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/249,216
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,210
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/226,681
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,759
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/225,213
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/227,182
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,214
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/235,836
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/230,438
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/215,135
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/225,266
```



```
; CURRENT APPLICATION NUMBER: US/10/769,507
; CURRENT FILING DATE: 2004-01-30
; PRIOR FILING DATE: 2004-02-28
; PRIOR FILING DATE: 2001-11-23
; PRIOR APPLICATION NUMBER: 09/288,037
; PRIOR FILING DATE: 1999-04-08
; PRIOR APPLICATION NUMBER: 60/081,037
; PRIOR FILING DATE: 1998-04-08
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 4
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-769-507-4

Alignment Scores:
Pred. No.: 1.9e-30      Length: 295
Score: 472.00          Matches: 100
Percent Similarity: 54.5%      Conservative: 52
Best Local Similarity: 35.8%    Mismatches: 101
Query Match: 10.7%            Indels: 26
DB: 5                      Gaps: 4

US-10-768-158-1 (1-2419) x US-10-769-507-4 (1-295)
Qy 90 GCGGTGGGTGCGCCCTTCTGCGCGGGAAGATGGAGGAGATCGCAACTTCCCGGTG 149
Db 17 GlyValProLeuIleTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla 36
Qy 150 CGGCCAGCAGCGTGTGATCGTACCTACCCCAAGTCCCGCACCAGCTGCTCCAGGAG 209
Db 37 ArgProAspLeuLeuIleAsnThrTyrProLysSerGlyThrTrpValSerGln 56
Qy 210 GTGTCTACTTGGTGGAGCGCGTACCCCGCATGAGATCGGCTTGATGAACATCGAC 269
Db 57 IleLeuAspMetIleTyrGlnGlyAspLeuGluLysCysAsnArgAlaProIleTyr 76
Qy 270 GAGCAGCTCCCGGTCTCGGAGTACCCACAGCGG-----GGCCTGGACATCATC 317
Db 77 ValArgValProPheLeuGluValAsnAspProGlyLuproSerGlyLeuGluThrLeu 96
Qy 318 AAGAACTGACCTCTCCCGCCTCATCAAGACCCACCTGCCCTACCGCTTCTCCCTCT 377
Db 97 LysAspThrProProArgLeuIleLysSerHisLeuProLeuAlaLeuLeuProGln 116
Qy 378 GACCTCCACAATGAGACTCCAAAGGTCTATATATGCTCGCAACCCCAAGGATCTGGT 437
Db 117 ThrLeuLeuAspGlnLysValLysValValValValValValValValValValVal 136
Qy 438 GTGTCTTATTATCAGTTCCACCGCTCTCTCCGACCACTAGCTACCGAGGACCTTTCAA 497
Db 137 ValSerTyrTyrHisPheHisArgMetGluLysAlaHisProGluProGlyThrTrpAsp 156
Qy 498 GAATTCCTCCGGAGGTTTATGAATGATATAGCTGGCTACCGCTCTCGTTTGACAGCTG 557
Db 157 SerPheLeuGluLysPheMetAlaGlyLysValSerTyrGlySerTyrGlnHisVal 176
Qy 558 CAGGAGTCTGGGAGCACCAGCATGCGTCCAAAGTCTTTTCTCAAGTATGAAGACATG 617
Db 177 GlnIuTrpGluLeuSerArgThrHisProValLeuTyrLeuPheTyrGluAspMet 196
Qy 618 CATCGGACCTGGTGGAGCGTGTGGAGCGTCCAGATTCCTGGGGGTGTCTCTGTGAC 677
Db 197 LysGluAsnProLysArgGluIleGlnLysIleLeuGluPheValGlyArgSerLeuPro 216
Qy 678 AAGCCCGCAGCTGGAAGCCCTGACGAGCATGCG-----710
Db 217 GluGluThrMetAspPheMetValGlnHisThrSerPheLysGluMetLysLysAsnPro 236
Qy 711 -----CACCAGCTGTGTGACGACGCTGCTCAACGCTGAGGCC 746
Db 237 MetThrAsnTyrThrThrValProGlnGluLeuMetAspHisSerIleSer----- 253

; CURRENT APPLICATION NUMBER: US/10/769,507
; CURRENT FILING DATE: 2004-01-30
; PRIOR FILING DATE: 2004-02-28
; PRIOR FILING DATE: 2001-11-23
; PRIOR APPLICATION NUMBER: 09/288,037
; PRIOR FILING DATE: 1999-04-08
; PRIOR APPLICATION NUMBER: 60/081,037
; PRIOR FILING DATE: 1998-04-08
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 4
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-769-507-4

Alignment Scores:
Pred. No.: 1.9e-30      Length: 295
Score: 472.00          Matches: 100
Percent Similarity: 54.5%      Conservative: 52
Best Local Similarity: 35.8%    Mismatches: 101
Query Match: 10.7%            Indels: 26
DB: 5                      Gaps: 4

US-10-768-158-1 (1-2419) x US-10-769-507-4 (1-295)
Qy 90 GCGGTGGGTGCGCCCTTCTGCGCGGGAAGATGGAGGAGATCGCAACTTCCCGGTG 149
Db 17 GlyValProLeuIleTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla 36
Qy 150 CGGCCAGCAGCGTGTGATCGTACCTACCCCAAGTCCCGCACCAGCTGCTCCAGGAG 209
Db 37 ArgProAspLeuLeuIleAsnThrTyrProLysSerGlyThrTrpValSerGln 56
Qy 210 GTGTCTACTTGGTGGAGCGCGTACCCCGCATGAGATCGGCTTGATGAACATCGAC 269
Db 57 IleLeuAspMetIleTyrGlnGlyAspLeuGluLysCysAsnArgAlaProIleTyr 76
Qy 270 GAGCAGCTCCCGGTCTCGGAGTACCCACAGCGG-----GGCCTGGACATCATC 317
Db 77 ValArgValProPheLeuGluValAsnAspProGlyLuproSerGlyLeuGluThrLeu 96
Qy 318 AAGAACTGACCTCTCCCGCCTCATCAAGACCCACCTGCCCTACCGCTTCTCCCTCT 377
Db 97 LysAspThrProProArgLeuIleLysSerHisLeuProLeuAlaLeuLeuProGln 116
Qy 378 GACCTCCACAATGAGACTCCAAAGGTCTATATATGCTCGCAACCCCAAGGATCTGGT 437
Db 117 ThrLeuLeuAspGlnLysValLysValValValValValValValValValValVal 136
Qy 438 GTGTCTTATTATCAGTTCCACCGCTCTCTCGGAGTACCCACAGCGG-----GGCCTGGACATCATC 497
Db 137 ValSerTyrTyrHisPheHisArgMetGluLysAlaHisProGluProGlyThrTrpAsp 156
Qy 498 GAATTCCTCCGGAGGTTTATGAATGATATAGCTGGCTACCGCTCTCGTTTGACAGCTG 557
Db 157 SerPheLeuGluLysPheMetAlaGlyLysValSerTyrGlySerTyrGlnHisVal 176
Qy 558 CAGGAGTCTGGGAGCACCAGCATGCGTCCAAAGTCTTTTCTCAAGTATGAAGACATG 617
Db 177 GlnIuTrpGluLeuSerArgThrHisProValLeuTyrLeuPheTyrGluAspMet 196
Qy 618 CATCGGACCTGGTGGAGCGTGTGGAGCGTCCAGATTCCTGGGGGTGTCTCTGTGAC 677
Db 197 LysGluAsnProLysArgGluIleGlnLysIleLeuGluPheValGlyArgSerLeuPro 216
Qy 678 AAGCCCGCAGCTGGAAGCCCTGACGAGCATGCG-----710
Db 217 GluGluThrMetAspPheMetValGlnHisThrSerPheLysGluMetLysLysAsnPro 236
Qy 711 -----CACCAGCTGTGTGACGACGCTGCTCAACGCTGAGGCC 746
Db 237 MetThrAsnTyrThrThrValProGlnGluLeuMetAspHisSerIleSer----- 253
```

Qy 558 CAGAGTTCTGGAGACCGCATGCGTCTTCTCAAGTATGAAGCATG 617
Db 177 GlnGluTrpTrpGluLeuSerArgThrHisProValLeuTyrLeuPheTyrGluAspMet 196
Qy 618 CATCGGACCTGGTGACGATGGTGAGCAGCTGCCAGATTCCTCGGGGGTCTCTGTGAC 677
Db 197 LysGluAsnProLysArgGluLeuGlnHisLeuGluPheValGlyArgSerLeuPro 216
Qy 678 AAGCCCGCAGCTGGAAGCCCTGACGAGCAGCTGC----- 710
Db 217 GluGluThrMetAspPheMetValGlnHisSerPheLysGluMetLysLysAsnPro 236
Qy 711 -----CACGAGCTGGTGGACCGAGCTGCTGCAACGCTGAGGCC 746
Db 237 MetThrAsnTyrThrValProGlnGluLeuMetAspHisSerIleSer----- 253
Qy 747 CTGCCC--GTGGCCGGGGAGAGTTGGGCTGTGGAGGACATCTTCACCGTCTCCATG 803
Db 254 ---ProPheMetArgLysGlyMetAlaGlyAspTrpLysThrPheThrValAlaGln 272
Qy 804 AATGAGAAGTTGACTGGTGTATATAACAGAGATGGGAAGTGTGACCTCACGTTT 860
Db 273 AsnGluArgPheAspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe 291

RESULT 11

US-10-072-012-666
; Sequence 666, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zernhosen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07

; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 666
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-072-012-666

Alignment Scores: 5.01e-30 Length: 296
Pred. No.: 467.00 Matches: 100
Score: 55.4% Conservative: 58
Percent Similarity: 35.1% Mismatches: 105
Best Local Similarity: 10.6% Indels: 22
Query Match: 4 Gaps: 5
DB: 4

US-10-768-158-1 (1-2419) x US-10-072-012-666 (1-296)

Qy 66 GAGAGCAAGTACTTCGAGTTCCATGGCGTGGCGCTCGCCCTTCTGCCGCGGAAGATG 125
Db 10 GlnThrLysLeuLysGluValAlaGlyLeuProLeuGlnAlaProThrValAsnTrp 29
Qy 126 GAGAGATCGCCAACTTCCTCCGGTGGCGGCCAGCGACGTGGATGTCACCTACCCCAAG 185
Db 30 SerGlnIleGlnThrPheLysAlaLysProAspAspLeuLeuLeuCysThrTyrProLys 49
Qy 186 TCCGGACACGAGTTCGTGGAGGAGGTGCTACTTGTGTAGCCAGGCGCTGACCCCGAT 245
Db 50 SerGlyThrThrTrpIleGlnGluIleValAspMetIleGluGlnAsnGlyAspValGlu 69
Qy 246 GAGATCGGCTTGATGAACATCGACGAGCTCCCGCTCTGGAGTAC-----CCA 296
Db 70 LysCysGlnArgThrIleIleGlnHisArgHisProPheIleGluTrpAlaArgProPro 89
Qy 297 CAGCCG---GGCTGGACATCATCAAGAACTGACCTCTCCCGCTCATCAAGAGCAC 353
Db 90 GlnProSerGlyValAspLysAlaAsnAlaMetProAlaProArgIleLeuArgThrHis 109
Qy 354 CTGCCCTACCGCTTCTGCGCTCTGACCTCCACAAATGGAGACTCCAAAGGTCTATATG 413
Db 110 LeuProThrGlnLeuLeuProProSerPheTrpThrAsnAsnCysLysPheLeuTyrVal 129
Qy 414 GCTCGCAACCCCAAGGATCTGGTGTCTTATATCATGTTCCACCGCTCTCTCGGACC 473
Db 130 AlaArgAsnAlaLysAspCysMetValSerTyrTyrHisPheTyrArgMetSerGlnVal 149
Qy 474 ATGAGCTACGAGCACCTTTCAGAAATTCCTCGCGGAGTTTATGATGATAGCTGGGC 533
Db 150 LeuProAspProGlyThrTrpAsnGluTyrPheGluThrPheIleAsnGlyLysValSer 169
Qy 534 TACGGCTCCTGGTTTGGACGCTGACGAGTCTCGGAGCACCGCATGGAGCTCGAAGCTG 593
Db 170 TrpGlySerTrpPheAspHisValLysGlyTrpTrpGluLeuLeuArgAspArgTyrGlnIle 189
Qy 594 CTTTTTCTCAAGTATGAAGACATGCATCGGACCTGGTGACGATGGTGGAGCAGCTGCC 653
Db 190 LeuPheLeuPheTyrGluAspValLysArgAspProLysArgGluIleGlnLysValMet 209
Qy 654 AGATTCTGGGGGTGCTC-----TGTCACAAAGCCCGAGCTGGAGCCCTG 698
Db 210 GlnPheMetGlyLysAsnLeuAspGluValValAspLysIleValLeuGluThrSer 229
Qy 699 ACGGAGCACTGCCACCACTGGTGGACAGTGTGCAACGCTGAGGCGCTGCC----- 752
Db 230 PheGluLysMetLysGlu-----AsnProMetThrAsnArgSerThrValProLysSer 247
Qy 753 -----GTGGCGCGGGAGAGTGTGGCTGTGGAAGAC 785

```
Db 248 ValLeuaspGlnSerIleSerProPheMetArgLysGlyThrValGlyAspTrpLysAsn 267
Qy 786 ATCTTCACCGTCTCCATGAATGAGAAGTTTGACATTGGTGATATAACAGCAAGATGGGAAG 845
Db 268 HisPheThrValAlaGlnAsnAspArgPheAspGluIleTyrLysGlnLysMetGlyGly 287
Qy 846 TGTACCTCACGTTT 860
Db 288 ThrSerLeuAsnPhe 292

RESULT 12
US-10-206-839-3
; Sequence 1, Application US/10206839
; Publication No. US20030099977A1
; GENERAL INFORMATION:
; APPLICANT: Guida, Marco
; APPLICANT: Kurth, Janice
; TITLE OF INVENTION: Genotyping Human Phenol Sulfotransferase
; TITLE OF INVENTION: (STP2)
; FILE REFERENCE: 4389-6 (Formerly SEQ-16P)
; CURRENT APPLICATION NUMBER: US/10/206,839
; CURRENT FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: 09/328,174
; PRIOR FILING DATE: 1999-06-08
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 295
; TYPE: PRT
; ORGANISM: H. sapiens
US-10-206-839-3

Alignment Scores:
Pred. No.: 4.22e-29 Length: 295
Score: 456.00 Matches: 94
Percent Similarity: 54.5% Conservative: 58
Best Local Similarity: 33.7% Mismatches: 101
Query Match: 10.4% Indels: 26
DB: 4 Gaps: 4

US-10-768-158-1 (1-2419) x US-10-206-839-3 (1-295)
Qy 90 GCGGTGCGGTCGCGCCCTTCGCGCGGGAAGATGGAGAGATCGCCAACTTCCCGTG 149
Db 17 GlyValProLeuIleLysTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla 36
Qy 150 CGGCCCGCAGCGTGTGGATCGTACCTACCCCAAGTCCGCGCAGCGTCTCTCGCAGG 209
Db 37 ArgProAspAspLeuLeuIleSerThrTyrProLysSerGlyThrTrpValSerGln 56
Qy 210 GTGCTCTACTTGTGAGCGCGGCTGACCCCGATGAGATCGGCTTGATGAACATCGAC 269
Db 57 IleLeuAspMetIleTyrGlnGlyGlyAspLeuGluLysCysHisArgAlaProIlePhe 76
Qy 270 GAGCAGCTCCCGTCTCGAGTACCCACAGCGG-----GGCTGGACATCATC 317
Db 77 MetArgValProPheLeuGluPheLysValProGlyIleProSerGlyMetGluThrLeu 96
Qy 318 RAGNACTGACTCTCCCGGCTCATCAAGAGCCACCTGCGCTACCGCTTTCGCCCTCT 377
Db 97 LysAsnThrProAlaProArgLeuLeuLysThrHisLeuProLeuAlaLeuLeuProGln 116
Qy 378 GACCTCCACAATGGAGACTCCAAAGTCTATATGCTGCGCAACCCCAAGATCTGGTG 437
Db 117 ThrLeuLeuAspGlnLysValValValTyrValAlaAlaArgAsnAlaLysAspValAla 136
Qy 438 GTGCTTATTATCAGTTCCACCGCTCTCTCGGACCATGAGCTACCGAGGACCTTTCAA 497
Db 137 ValSerTyrTyrHisPheTyrHisMetAlaLysValTyrProHisProGlyThrTrpGlu 156
Qy 498 GAATCTCCCGGAGTTTATGAATGATAAGCTGGCTACGCTCTCTGTTTGAGCAGCTG 557
Db 157 SerPheLeuGluLysPheMetAlaGlyGluValSerTyrGlySerTyrTrpTyrGlnHisVal 176
```

```
Qy 558 CAGAGATTCTGGGAGACCCGATCGAATCGAATCTGCTCTTTTCTCAAGTATGAAGACATG 617
Db 177 GlnGluTrpTrpGluLeuSerArgThrHisProValLeuTyrLeuPheTyrGluAspMet 196
Qy 618 CATCGGACCTGGTGACGATGGTGAGCAGCTGCCAGATTCCTCGGGGTCTCTGTGAC 677
Db 197 LysGluAsnProLysArgGluIleGlnLysIleLeuGluPheValGlyArgSerLeuPro 216
Qy 678 AAGCCCGCAGCTGGAAGCCCTGACGAGCAGCTGC----- 710
Db 217 GluGluThrValAspLeuMetValGluHisThrSerPheLysGluMetLysLysAsnPro 236
Qy 711 -----CACAGCTGTGTGGACCAAGTCTGCAACGCTGAGGCC 746
Db 237 MetThrAsnTyrThrThrValArgGluPheMetAspHisSerIleSer----- 253
Qy 747 CTGCCC---GTGGCGCGGGAAGATTGGGCTGTGGGAAGGACATCTTCACCGTCTCCATG 803
Db 254 ---ProPheMetArgLysGlyMetAlaGlyAspTrpLysThrThrPheThrValAlaGln 272
Qy 804 AATGAGAAGTTTGACTTGGTGTATATAACAGAAAGATGGAAAGTGTGACCTCACGTTT 860
Db 273 AsnGluArgPheAspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe 291

RESULT 13
US-10-769-507-1
; Sequence 1, Application US/10769507
; Publication No. US20050019788A1
; GENERAL INFORMATION:
; APPLICANT: Squires, James E.
; APPLICANT: Lin, Zhihong E.
; TITLE OF INVENTION: Genetic Markers for Skatole Metabolism
; FILE REFERENCE: P06153US03
; CURRENT APPLICATION NUMBER: US/10/769,507
; CURRENT FILING DATE: 2004-01-30
; PRIOR APPLICATION NUMBER: 10/024,628
; PRIOR FILING DATE: 2001-11-23
; PRIOR APPLICATION NUMBER: 09/288,037
; PRIOR FILING DATE: 1999-04-08
; PRIOR APPLICATION NUMBER: 60/081,037
; PRIOR FILING DATE: 1998-04-08
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 295
; TYPE: PRT
; ORGANISM: porcine
US-10-769-507-1

Alignment Scores:
Pred. No.: 4.22e-29 Length: 295
Score: 456.00 Matches: 98
Percent Similarity: 54.2% Conservative: 51
Best Local Similarity: 35.6% Mismatches: 108
Query Match: 10.4% Indels: 18
DB: 5 Gaps: 3

US-10-768-158-1 (1-2419) x US-10-769-507-1 (1-295)
Qy 90 GCGGTGCGGTCGCGCCCTTCGCGCGGGAAGATGGAGAGATCGCCAACTTCCCGTG 149
Db 17 GlyValProLeuIleLysTyrPheAlaGluAlaLeuGlyProLeuGluSerPheGlnAla 36
Qy 150 CGGCCCGCAGCGTGTGGATCGTACCTACCCCAAGTCCGCGCAGCGTCTCTCGCAGG 209
Db 37 TrpProAspAspValLeuIleSerThrTyrProLysSerGlyThrTrpValSerGlu 56
Qy 210 GTGCTCTACTTGTGAGCGCGGCTGACCCCGATGAGATCGGCTTGATGAACATCGAC 269
Db 57 IleLeuAspLeuIleTyrGlnGlyGlyAspLeuGlnLysCysGlnArgAlaProIlePhe 76
Qy 270 GAGCAGCTCCCGGCTCTCGAGTACCCACAGCGG-----GGCTGGACATCATC 317
```



```
QY 150 CGGCCACGACGCTGTGTGATCGTACCTACCCCAAGTCCGACACGAGCTTCTCGAGGAG 209
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 37 ArgProAspLeuLeuIleSerThrTyrProLysSerGlyThrTrpValSerGln 56
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 210 GTGGTCTACTTGGTACGACGGCGCTGACCCGATGAGATCGGCTTGATGAACATCGAC 269
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 57 IleLeuAspMetIleTyrGlnGlyAspLeuGluLysCysHisArgAlaProIlePhe 76
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 270 GAGCAGCTCCCGGTCTCGAGTACCCACAGCCG-----GGCCTGGACATCATC 317
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 77 MetArgValProPheLeuGluPheLysAlaProGlyIleProSerGlyMetGluThrLeu 96
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 318 AAGGAACGTACCTCTCCCGCTCATCAAGAGCCACCTGCGCTTCTTCCCTCT 377
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 97 LysAspThrProAlaProArgLeuLysThrHisLeuProLeuAlaLeuLeuProGln 116
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 378 GACTCTCCCAATGGAGATCCCAAGTCTATATGCTCGCTCCCAAGTCTGGTG 437
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 117 ThrLeuLeuAspGlnLysValValTyrValAlaAlaArgAsnAlaLysAspValAla 136
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 438 GTGTCTTATTATCAAGTCTCCACCGCTCTCTCGCGACCATGAGCTACCGAGGACCTTTCAA 497
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 137 ValSerTyrTyrHisPheTyrHisMetAlaLysValHisProGluProGlyThrTrpAsp 156
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 498 GAATTTCTCCGGAGGTTTATGAATGATAAGTGGGCTACCGCTCTCGTCTTGTGACACGTG 557
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 257 ArgLysGlyMetAlaGlyAspTrpLysThrPheThrValAlaGlnAsnGluArgPhe 276
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 816 GACTTGGTGATATAACAGAAAGTGGAAAGTGGACCTCAGCTT 860
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 277 AspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe 291
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

RESULT 19

```
US-10-734-049A-245
; Sequence 245, Application US/10734049A
; Publication No. US20050042624A1
; GENERAL INFORMATION:
; APPLICANT: ITOH, Kiyoko
; APPLICANT: SHICHIO, Shigeki
; TITLE OF INVENTION: TUMOR ANTIGEN
; FILE REFERENCE: Q-78382
; CURRENT APPLICATION NUMBER: US/10/734,049A
; PRIORITY FILING DATE: 2003-12-12
; PRIORITY FILING DATE: PCT/JP02/05799
; PRIORITY FILING DATE: 2002-06-11
; PRIORITY FILING DATE: JP 2001/177058
; PRIORITY FILING DATE: 2001-06-12
; PRIORITY FILING DATE: JP 2001/250728
; PRIORITY FILING DATE: 2001-08-21
; NUMBER OF SEQ ID NOS: 408
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 245
; LENGTH: 295
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-734-049A-245

Alignment Scores: 1.11e-28 Length: 295
Pred. No.: 451.00 Matches: 93
Score: 54.5% Conservative: 57
Best Local Similarity: 33.8% Mismatches: 107
Query Match: 10.2% Indels: 18
DB: 5 Gaps: 3

US-10-768-158-1 (1-2419) x US-10-734-049A-245 (1-295)
QY 90 GGCCTGGCGGTGCGCGCTTCTGCGCGGGAAGATGAGGAGATCGCCAACTTCCCGGTG 149
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 17 GlyValProLeuIleLysTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla 36
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 150 CGGCCACGACGCTGTGTGATCGTACCTACCCCAAGTCCGCGACACGCTTCTCGAGGAG 209
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 37 ArgProAspLeuLeuIleSerThrTyrProLysSerGlyThrTrpValSerGln 56
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 210 GTGTCTACTTGGTACGACGGCGCTGACCCCGATGAGATCGGCTTGATGAACATCGAC 269
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 57 IleLeuAspMetIleTyrGlnGlyAspLeuGluLysCysHisArgAlaProIlePhe 76
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 270 GAGCAGCTCCCGGTGCTCGGAGTACCCACAGCCG-----GGCCTGGACATCATC 317
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 77 MetArgValProPheLeuGluPheLysAlaProGlyIleProSerGlyMetGluThrLeu 96
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 318 AAGGAACGTACCTCTCCCGCTCATCAAGAGCCACCTGCGCTTCTTCCCTCT 377
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 97 LysAspThrProAlaProArgLeuLysThrHisLeuProLeuAlaLeuLeuProGln 116
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 378 GACTCTCCCAATGGAGATCCCAAGTCTATATGCTCGCTCCCAAGTCTGGTG 437
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 117 ThrLeuLeuAspGlnLysValValTyrValAlaAlaArgAsnAlaLysAspValAla 136
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 438 GTGTCTTATTATCAAGTCTCCACCGCTCTCTCGCGACCATGAGCTACCGAGGACCTTTCAA 497
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 137 ValSerTyrTyrHisPheTyrHisMetAlaLysValHisProGluProGlyThrTrpAsp 156
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 498 GAATTTCTCCGGAGGTTTATGAATGATAAGTGGGCTACCGCTCTCGTCTTGTGACACGTG 557
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 257 ArgLysGlyMetAlaGlyAspTrpLysThrPheThrValAlaGlnAsnGluArgPhe 276
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 816 GACTTGGTGATATAACAGAAAGTGGAAAGTGGACCTCAGCTT 860
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 277 AspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe 291
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 20
US-10-072-012-586
; Sequence 586, Application US/10072012
```



```
Db 94 LeuProLeuGluLeuLeuProLysSerPheLeuSerSerLysAlaLysIleIleTyrVal 113
Qy 414 GCTCGCAACCCCAAGGATCTGGTGTCTTATTATCAGTTCCACCGCTCTCTCGGACC 473
Db 114 LeuArgAsnProLysAspValAlaValSerTyrTyrHisPheSerArgSerHisLysAsp 133
Qy 474 ATG---AGCTACCGAGGACCTTTCAAGAAATTCGCCGAGGTTTATGAATGATAAGCTG 530
Db 134 LeuProAlaAspProGlyThrPheGluGluPheLeuGluAlaPheLeuAsnGlyLysVal 153
Qy 531 GGCTACGCTCCTCGTTTGACACGTCGACGAGTTCGGGAGCACCGCATGGACTCGAAC 590
Db 154 LeuTyrGlySerTyrPheAspHisValLeuGlyTyrPheGluLeuArgProGluProGln 173
Qy 591 GTGCTTTTTCATAGTATGAAGACATGCATCGGACCTGTGGTCATGGTGGAGCAGCTG 650
Db 174 ValLeuPheLeuAspTyrGluAspLeuLysGluAspProAlaGlyGluIleLysLysIle 193
Qy 651 GCCAGATCTCTGGGGTGTCTCTGTGACAAGCCCGAGCTGGAAGCCCTCGACGACACTGC 710
Db 194 AlaGluPheLeuGlyLeuProLeuSerGluGluLeuAspLysLeuLeuAspHisSer 213
Qy 711 CAC-----CAGCTGGTGACCATGCTGTCACGCTGAGGCCCTGCCCGTGGC--- 758
Db 214 SerPhePheLeuMetLysLeuAsnProLeuSerAsnTyrGluThrLeuCysLeuGlyLys 233
Qy 759 -----CGGGGAAGAGTTGGCTGTGAAGGACATC 788
Db 234 SerLysGlyArgLysSerProPheMetArgLysGlyLeuValGlyAspTrpLysAsnTyr 253
Qy 789 TTCACCTCTCCATGAATGAGAAGTTTGACTTGGTGTGATTAACAGAGAAG 836
Db 254 PheThrProGluGlnAsnGluLysPheAspLysValIleLysGluLys 269
```

RESULT 24

```
US-10-072-012-867
; Sequence 867, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zernusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31

Alignment Scores:
Pred. No.: 4,17e-28 Length: 269
Score: 444.00 Matches: 104
Percent Similarity: 55.9% Conservative: 39
Best Local Similarity: 40.6% Mismatches: 83
Query Match: 10.1% Indels: 30
DB: 4 Gaps: 7

US-10-768-158-1 (1-2419) x US-10-072-012-867 (1-269)
Qy 141 TTCCCGGTGGCGCCACGACGCTGTGTGATCGTACCTACCCCAAGTCCGGACAGCTTG 200
Db 20 PheGlnAlaArgProAspAspValLeuIleAlaGlyTyrProLysSerGlyThrTrp 39
Qy 201 CTCGAGGAGGTGTCTACTTGGTGAGCCAGCGGCTGCAC-----CCCGAT 245
Db 40 LeuGlnGluIleLeuSerLeuHisProAsnValGlyAspPheGluProSerProSerAsp 59
Qy 246 GAGATCGGCTTGATGAACATCGACGACGCTCCCGTCTCTGGAGTAGTACCCACAGCCGGC 305
Db 60 ProLeuLeuPheArgAsn-----ProTrpLeuGluTyrProLys---Gly 73
Qy 306 CTGGACATCATCAAGGAACGTG-----ACCTCTCCCGCCCTCATCAAGAGCCAC 353
Db 74 GluAspTrpTyrGluThrLeuLysProMetProSerSerProArgLeuIleLysThrHis 93
Qy 354 CTGCCCTACCGCTTCTCGCCCTCTGACCTCCACATGGAGACTCCCAAGGTCTATATG 413
Db 94 LeuProLeuGluLeuLeuProLysSerPheLeuSerSerLysAlaLysIleIleTyrVal 113
Qy 414 GCTCGCAACCCCAAGGATCTGGTGTCTTATTATCAGTTCCACCGCTCTCTCGGACC 473
Db 114 LeuArgAsnProLysAspValAlaValSerTyrTyrHisPheSerArgSerHisLysAsp 133
Qy 474 ATG---AGCTACCGAGGACCTTTCAAGAAATTCGCCGAGGTTTATGAATGATAAGCTG 530
Db 134 LeuProAlaAspProGlyThrPheGluGluPheLeuGluAlaPheLeuAsnGlyLysVal 153
Qy 531 GGCTACGCTCCTGGTTTGACGAGTTCGGGAGCACCGCATGGACTCGAAC 590
Db 154 LeuTyrGlySerTyrPheAspHisValLeuGlyTyrPheGluLeuArgProGluProGln 173
Qy 591 GTGCTTTTTCATAGTATGAAGACATGCATCGGACCTGTGGAGTGTGGAGCAGCTG 650
Db 174 ValLeuPheLeuAspTyrGluAspLeuLysGluAspProAlaGlyGluIleLysLysIle 193
Qy 651 GCCAGATCTCTGGGGTGTCTCTGTGACAAGCCCGAGCTGGAAGCCCTCGACGACACTGC 710
```



```

Db      67 ValGluLysCysLysGluAspAlaIlePheAsnArgIleProTyrLeuGluCysArgAsn 86
      300 CCG-----GGCTGCATCATCAAGAACTGACTCTCCCGCGCTCATCAAG 347
      87 GluAspLeuIleAsnGlyIleLysGlnLeuLysGluLysGluSerProArgIleVallys 106
      348 AGCCACCTGCGCTACCGCTTCTGCGCTCTGACCTCCCAATGAGAGATCCCAAGGTATC 407
      107 ThrHisLeuProLysValLeuProAlaSerPheThrGluLysAsnCysLysMetile 126
      408 TATATGGCTCCCAACCCCAAGACTGCTGCTGCTTATTATCATGATCCACCGCTCTCTG 467
      127 TyrLeuCysArgAsnAlaLysAspValIleValSerTyrTyrPheLeuLeuMetile 146
      468 CGGACCATGAGTACCGAGGACCTTTCAAGAAATTTCTCCCGGAGTTTATGAATGATAAG 527
      147 ThrSerTyrProAsnProLysSerPheSerGluPheValGluLysPheMetGlnGlyGln 166
      528 CTGGGCTACGCTCTCTGTTGAGACGTGAGGAGTTCTGGAGACCGCATGAGTCTG 587
      167 ValProTyrGlySerTrpTyrAspHisValLysAlaTrpTrpGluLysSerLysAsnSer 186
      588 AACGTGCTTTTCTCAAGTATGAACATGATCGGACCTGTCGACGATCGTGGAGCAG 647
      187 ArgValLeuPheMetPheTyrGluAspMetLysGluAspIleArgGluValVallys 206
      648 CTGGCCAGATTCCTGGGGGTCTCTGTGACAAAGCCAGCTGGAAGCCCTGACCGAGCAC 707
      207 LeuIleGluPheLeuGluArgLysProSerAlaGluLeuValAspArgIleGlnHis 226
      708 TGCCAC-----CAGCTGGTGGACGAC 728
      227 ThrSerPheGlnGluMetLysAsnAsnProSerThrAsnTyrThrMetProGluGlu 246
      729 TGCTGCAAGCTGAGGCCCTGCCCC---GTGGCCGGGGAAGTGGCTCTGGAAGCAC 785
      247 MetMetAsnGlnLysValSerProPheMetArgLysGlyIleIleGlyAspTrpLysAsn 266
      786 ATCTTTCACCGTCTCCATGAATGAGAAGTTTGACTTGGTGTATAACAGAAAGTGGAAAG 845
      267 HisPheProGluAlaLeuArgGluArgPheAspGluHisTyrLysGlnGlnMetLysAsp 286
      846 TGTGACCTCAGCTTT 860
      287 CysThrValLysPhe 291

RESULT 27
US-10-201-525-13
; Sequence 13, Application US/10201525
; Publication No. US20060009631A1
; GENERAL INFORMATION:
; APPLICANT: Board of Regents of the University of Oklahoma
; TITLE OF INVENTION: TYROSYLPROTEIN SULFOTRANSFERASES AND METHODS OF USE THEREOF
; FILE REFERENCE: 5827.005
; CURRENT APPLICATION NUMBER: US/10/201.525
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/785,343
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: PCT/US99/16750
; PRIOR FILING DATE: 1999-07-23
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Mus musculus
; US-10-201-525-13

Alignment Scores:
Pred. No.: 6,5e-27 Length: 295
Score: 430.00 Matches: 94
Percent Similarity: 53.0% Conservative: 57

```

```

Best Local Similarity: 33.0% Mismatches: 116
Query Match: 9.8% Indels: 18
DB: 5 Gaps: 3

US-10-768-158-1 (1-2419) x US-10-201-525-13 (1-295)

QY      60 GAGTTCGAGACCAAGTACTTCGAGTTCATGGCGTGGCGTCCGCTTCTGCGCGGG 119
DB      7 GluTyrTyrGluValPheGlyGluPheArgGlyValLeuMetAspLysArgPheThrLys 26
      120 AAGATGAGGAGATCGCCAACTTCCCGTGGCGGCCAGCGAGCGTGTGATCGTCACTTAC 179
DB      27 TyrTrpGluAspValGluMetPheLeuAlaArgProAspAspLeuValIleAlaThrTyr 46
      180 CCCAAGTCCGACACCGAGTCTGCTCAGGAGGTGTCTACTTGGTGAGCCAGGCGCTGAC 239
DB      47 ProLysSerGlyThrThrTrpIleSerGluValValTyrMetIleTyrLysGluGlyAsp 66
      240 CCCGATGAGATCGGCTTGATGAACATCAGCAGCAGCTCCCGTCTCGGAGTATCCACAG 299
DB      67 ValGluLysCysLysGluAspAlaIlePheAsnArgIleProTyrLeuGluCysArgAsn 86
      300 CCG-----GGCTGGACATCATCAAGAACTGACTCTCCCGCGCTCATCAAG 347
DB      87 GluAspLeuIleAsnGlyIleLysGlnLeuLysGluLysGluSerProArgIleVallys 106
      348 AGCCACCTGCGCTACCGCTTCTGCGCTCTGACCTCCACAAATGGAGACTCCAAGGTATC 407
DB      107 ThrHisLeuProLysValLeuProAlaSerPheTrpGluLysAsnCysLysMetile 126
      408 TATATGGCTCGCAACCCCAAGGATCTGCTGCTGCTTATTATCATGATCCACCGCTCTCTG 467
DB      127 TyrLeuCysArgAsnAlaLysAspValAlaValSerTyrTyrPheLeuLeuMetile 146
      468 CGGACCATGAGTACCGAGGACCTTTCAAGAAATTTCTGCGGAGTTTATGAATGATAAG 527
DB      147 ThrSerTyrProAsnProLysSerPheSerGluPheValGluLysPheMetGlnGlyGln 166
      528 CTGGGCTACGCTCTCTGTTGAGCAGCTGCGAGGATTTCTGGAGACCGCATGAGTCTG 587
DB      167 ValProTyrGlySerTrpTyrAspHisValLysAlaTrpTrpGluLysSerLysAsnSer 186
      588 AACGTGCTTTTCTCAAGTATGAACATGATCGGACCTGTCGACGATCGTGGAGCAG 647
DB      187 ArgValLeuPheMetPheTyrGluAspMetLysGluAspIleArgGluValVallys 206
      648 CTGGCCAGATTCCTGGGGGTCTCTGTGACAAAGCCAGCTGGAAGCCCTGACCGAGCAC 707
DB      207 LeuIleGluPheLeuGluArgLysProSerAlaGluLeuValAspArgIleGlnHis 226
      708 TGCCAC-----CAGCTGGTGGACGAC 728
DB      227 ThrSerPheGlnGluMetLysAsnAsnProSerThrAsnTyrThrMetProGluGlu 246
      729 TGCTGCAAGCTGAGGCCCTGCCCC---GTGGCCGGGGAAGTGGCTCTGGAAGCAC 785
DB      247 MetMetAsnGlnLysValSerProPheMetArgLysGlyIleIleGlyAspTrpLysAsn 266
      786 ATCTTTCACCGTCTCCATGAATGAGAAGTTTGACTTGGTGTATAACAGAAAGTGGAAAG 845
DB      267 HisPheProGluAlaLeuArgGluArgPheAspGluHisTyrLysGlnGlnMetLysAsp 286
      846 TGTGACCTCAGCTTT 860
DB      287 CysThrValLysPhe 291

RESULT 28
US-10-072-012-585
; Sequence 585, Application US/10072012
; Publication No. US2004003493A1
; GENERAL INFORMATION:
; APPLICANT: Tchevnev, Velizar
; APPLICANT: Spytek, Kimberly

```


US-10-199-330-5

Alignment Scores:

Pred. No.: 3.1e-26 Length: 304
 Score: 422.00 Matches: 97
 Percent Similarity: 52.0% Conservative: 59
 Best Local Similarity: 32.3% Mismatches: 124
 Query Match: 9.6% Indels: 20
 DB: 4 Gaps: 5

US-10-768-158-1 (1-2419) x US-10-199-330-5 (1-304)

```

QY 21 ATGCGGAGAGCGAGCGGAGACCCCGGAGCGGAGGAGTTCGAG---AGCAAGTAC 77
D 1 MetAlaLysIleGluLysAsnAlaProThrMetGluLysLysProGluLeuPheAsnIle 20
QY 78 TTCGAGTTCCATCGCGTGGCGGCTGCGGCCCTCTCGCGCGGGAAGATGGAGGATCGCC 137
D 21 MetGluValAspGlyValProThrLeuIleLeuSerLysGluTrpTrpGluLysValCys 40
QY 138 AACTTCCGGTGGCGGCGGAGCGAGTGTGGATCTCACTACCCCAAGTCCGGCAGCAGC 197
D 41 AsnPheGlnAlaLysProAspAspLeuIleLeuAlaThrTyrProLysSerGlyThrThr 60
QY 198 TTGCTGCAGGAGTGGTCTACTTGGTGAGCGAGCGGCGCTGACCCCGATGAGATCGGCTTG 257
D 61 TrpMetHisGluIleLeuAspMetIleLeuAsnAspGlyAspValGluLysCysLysArg 80
QY 258 ATGAACATCGACGAGCGAGTCCCGGCTCTGGAG---TACCCA-----CAGCCG 302
D 81 AlaGlnThrLeuAspArgHisAlaPheLeuGluLeuLysPheProHisLysGluLysPro 100
QY 303 GGCTTGACATCATCAAGGAACCTCTCCCGCTCATCAAGAGCCACTCCCGCTAC 362
D 101 AspLeuGluPheValLeuGluMetSerSerProGlnLeuIleLysThrHisLeuProSer 120
QY 363 CGCTTTCTGCGCTGACCTCCACCAATGGAGACTCCCAAGTCACTATATGGCTCGCAAC 422
D 121 HisLeuIleProProSerIleTrpLysGluAsnCysLysIleValTyrValAlaArgAsn 140
QY 423 CCCAAGGATCTGGTGTCTTATATACAGTTCCACCGCTCTCTCGCGGACCATGAGTAC 482
D 141 ProLysAspCysLeuValSerTyrTyrHisPheHisArgMetAlaSerPheMetProAsp 160
QY 483 CGAGGCACCTTTCAGAAATTTGCGGAGGTTTATGAATGATAAGCTGGGCTACGGCTCC 542
D 161 ProGlnAsnLeuGluGluPheTyrGluLysPheMetSerGlyLysValValGlySer 180
QY 543 TGGTTTGACGACGTGCGAGGAGTCTGGGAGCACCGCATGAGCTCGAAGCTGCTTTTCTC 602
D 181 TrpPheAspHisValLysGlyTrpTrpAlaAlaLysAspMetHisArgIleLeuTyrLeu 200
QY 603 AAGTATCAAGACATGATCGGAGCTGGTGCAGTGTGGAGCAGCTGGCCAGATTCCTG 662
D 201 PheTyrGluAspIleLysLysAspProLysArgGluIleGluLysIleLeuLysPheLeu 220
QY 663 GGGGTGCTGTGACAGCCCGACCTGGAAGCCCTGACGAGCAGCTGCCACCACTGGT 722
D 221 GluLysAspIleSerGluGluIleLeuAsnLysIleIleTyrHisThrSerPheAspVal 240
QY 723 -----GACCATGCTGCAAGCTGAGCGCTGCGGCCCTGCGCGGGC----- 758
D 241 MetLysGlnAsnProMetThrAsnTyrThrThrLeuProThrSerIleMetAspHisSer 260
QY 759 -----CGGGAAGAGTTGGGCTGTGGAGGACATCTTCACCGTCTCC 800
D 261 IleSerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAla 280
QY 801 ATGAATGAGAAGTTTGACTTGGTGTATATAAACAAGAGATGGAAAGTGTGACCTCACCGTTT 860
D 281 GlnAsnGluGluPheAspLysAspTyrGlnLysLysMetAlaGlySerThrLeuThrPhe 300

```

RESULT 30

US-10-199-334-5

; Sequence 5, Application US/10199334
 ; Publication No. US2003008354A1
 ; GENERAL INFORMATION:
 ; APPLICANT: WOODAGE, Trevor et al.
 ; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
 ; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
 ; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
 ; FILE REFERENCE: CLO00669DIV-3
 ; CURRENT APPLICATION NUMBER: US/10199,334
 ; CURRENT FILING DATE: 2002-07-22
 ; PRIOR FILING DATE: 09/609,816
 ; PRIOR FILING DATE: 2001-03-27
 ; PRIOR APPLICATION NUMBER: 60/192,408
 ; PRIOR FILING DATE: 2000-03-27
 ; PRIOR APPLICATION NUMBER: 60/212,725
 ; PRIOR FILING DATE: 2000-06-20
 ; PRIOR APPLICATION NUMBER: 09/609,816
 ; PRIOR FILING DATE: 2000-07-03
 ; NUMBER OF SEQ ID NOS: 23
 ; SOFTWARE: Fast-Seq for Windows Version 4.0
 ; SEQ ID NO 5
 ; LENGTH: 304
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-199-334-5

Alignment Scores:

Pred. No.: 3.1e-26 Length: 304
 Score: 422.00 Matches: 97
 Percent Similarity: 52.0% Conservative: 59
 Best Local Similarity: 32.3% Mismatches: 124
 Query Match: 9.6% Indels: 20
 DB: 4 Gaps: 5

US-10-768-158-1 (1-2419) x US-10-199-334-5 (1-304)

```

QY 21 ATGCGGAGAGCGAGCGGAGACCCCGGAGCGGAGTTCGAG---AGCAAGTAC 77
D 1 MetAlaLysIleGluLysAsnAlaProThrMetGluLysLysProGluLeuPheAsnIle 20
QY 78 TTCGAGTTCCATCGCGTGGCGGCTGCGGCCCTCTCGCGCGGGAAGATGGAGGATCGCC 137
D 21 MetGluValAspGlyValProThrLeuIleLeuSerLysGluTrpTrpGluLysValCys 40
QY 138 AACTTCCGGTGGCGGCGGAGCGAGTGTGGATCTCACTACCCCAAGTCCGGCAGCAGC 197
D 41 AsnPheGlnAlaLysProAspAspLeuIleLeuAlaThrTyrProLysSerGlyThrThr 60
QY 198 TTGCTGCAGGAGTGGTCTACTTGGTGAGCGGCGCTGACCCCGATGAGATCGGCTTG 257
D 61 TrpMetHisGluIleLeuAspMetIleLeuAsnAspGlyAspValGluLysCysLysArg 80
QY 258 ATGAACATCGACGAGCGAGTCCCGGCTCTGGAG---TACCCA-----CAGCCG 302
D 81 AlaGlnThrLeuAspArgHisAlaPheLeuGluLeuLysPheProHisLysGluLysPro 100
QY 303 GGCTTGACATCATCAAGGAACCTGACCTCTCCCGCTCATCAAGAGCCACTCCCGCTAC 362
D 101 AspLeuGluPheValLeuGluMetSerSerProGlnLeuIleLysThrHisLeuProSer 120
QY 363 CGCTTTCTGCGCTGACCTCCCAATGGAGACTCCCAAGTCACTATATGGCTCGCAAC 422
D 121 HisLeuIleProProSerIleTrpLysGluAsnCysLysIleValTyrValAlaArgAsn 140
QY 423 CCCAAGGATCTGGTGTCTTATATACAGTTCCACCGCTCTCTCGCGGACCATGAGTAC 482
D 141 ProLysAspCysLeuValSerTyrTyrHisPheHisArgMetAlaSerPheMetProAsp 160
QY 483 CGAGGCACCTTTCAGAAATTTGCGGAGGTTTATGAATGATAAGCTGGGCTACGGCTCC 542
D 161 ProGlnAsnLeuGluGluPheTyrGluLysPheMetSerGlyLysValValGlySer 180

```



```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-108-875-5

Alignment Scores:
Pred. No.: 3,1e-26 Length: 304
Score: 422.00 Matches: 97
Percent Similarity: 52.0% Conservative: 59
Best Local Similarity: 32.3% Mismatches: 124
Query Match: 9.6% Indels: 20
DB: 6 Gaps: 5

US-10-768-158-1 (1-2419) x US-11-108-875-5 (1-304)
Qy 21 ATGCGGAGAGCGGCGGAGACCCAGCACCCCGGGGAGTTCGAG---AGCAAGTAC 77
Db 1 MetAlaLysIleGluLysAsnAlaProThrMetGluLysLysProGluLeuPheAsnIle 20
Qy 78 TTCAGATTCCATGCGGTGCGGCTCGGCCCTTCGCCGGGGAAGATGGAGGATCGCC 137
Db 21 MetGluValAspGlyValProThrLeuIleLeuSerLysGluTrpGluLysValCys 40
Qy 138 AACTTCCGGTGGCGGCGGAGCGTGTGGATCGTCACTACCCCAAGTCCGGCAGCAGC 197
Db 41 AsnPheGlnAlaLysProAspAspLeuLeuAlaThrTyrProLysSerGlyThr 60
Qy 198 TTGCTGCAGGAGGTGTCTACTTGTGTAGCGGCGTGAACCCCGATGAGATCGGCTTG 257
Db 61 TrpMetHisGluIleLeuAspMetIleLeuAsnAspGlyAspValGluLysCysLysArg 80
Qy 258 ATGAACATCGACGAGCTCCCGTCTGTGAG-----TACCCA-----CAGCGC 302
Db 81 AlaGlnThrLeuAspArgHisAlaPheLeuGluLysPheProHisLysLysGluLysPro 100
Qy 303 GGCTGTGACATCATCAGGAAGTCACTCTCCCGCTCATCAGAGCCAGCCCTCCCTAC 362
Db 101 AspLeuGluPheValLeuLysAspSerProGlnLeuIleLysThrHisLeuProSer 120
Qy 363 CGCTTTCTGCGCTCTGACCTCCCAATGGAGACTCCCAAGGTCACTATATGGCTCGCAAC 422
Db 121 HisLeuIleProProSerIleTrpLysGluAsnCysLysIleValTyrValAlaArgAsn 140
Qy 423 CCAAGATCTGGGTGTCTTATATCAGTTCCACCGCTCTCTGCGGACCATCAGGTAC 482
Db 141 ProLysAspCysLeuValSerTyrTyrHisPheHisArgMetAlaSerPheMetProAsp 160
Qy 483 CGAGGCACCTTTCAGAAATTCGCGGAGGTTTATGAATAGTACCTGGGTACGCTCC 542
Db 161 ProGlnAsnLeuGluGluPheTyrGluLysPheMetSerGlyLysValValGlyLys 180
Qy 543 TGGTTGAGCAGCTGCAGGAGTTCCTGGGAGCACCGCATGGACTCGAAGCTGCTTTTCTC 602
Db 181 TrpPheAspHisValLysGlyTrpTrpAlaAlaLysAspMetHisArgIleLeuTyrLeu 200
Qy 603 AAGTATGAAGACATGCATCGGACCTGGTGACGATGGTGGAGCAGTGGCCAGATCTCTG 662
Db 201 PheTyrGluAspIleLysLysAspProLysArgGluIleGluLysIleLeuLysPheLeu 220
Qy 663 GGGGTGCTCTGTACAGGCCACCTGGAGCCCTGACGAGCACTGCCACCAGCTGGTG 722
Db 221 GluLysAspLeuSerGluGluIleLeuAsnLysIleTyrHisThrSerPheAspVal 240
Qy 723 -----GACCAGTGTGCAAGCTCGAGGCCCTGCGCGTGGGC----- 758
Db 241 MetLysGlnAsnProMetThrAsnTyrThrLeuProThrSerIleMetAspHisSer 260
Qy 759 -----CGGGGAAGATTGGGTGTGGAGGACATCTTCACCGTCTCC 800
Db 261 IleSerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAla 280
Qy 801 ATGAATGAAGATTGACTTGGTGTATAAACAGAGATGGGAAGATGTGACCTCACCTTT 860
Db 281 GlnAsnGluGluPheAspLysAspTyrGlnLysLysMetAlaGlySerThrLeuThrPhe 300
```

```

RESULT 33
US-09-898-570-40
; Sequence 40, Application US/09898570
; Patent No. US20020123612A1
; GENERAL INFORMATION:
; APPLICANT: GERLACH, VALERIE L.
; APPLICANT: ELLERMAN, KAREN
; APPLICANT: MACDOUGALL, JOHN R.
; APPLICANT: SMITHSON, GLENDA
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND
; FILE REFERENCE: 15966-776CIP
; CURRENT APPLICATION NUMBER: US/09/898,570
; CURRENT FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 60/198,293
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: 60/198,645
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: 60/210,809
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/199,476
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/200,025
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/224,610
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: 60/200,024
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/199,880
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/218,591
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 60/271,814
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/215,855
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 09/839,446
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-898-570-40

Alignment Scores:
Pred. No.: 4,9e-26 Length: 283
Score: 419.50 Matches: 91
Percent Similarity: 53.0% Conservative: 57
Best Local Similarity: 32.6% Mismatches: 112
Query Match: 9.5% Indels: 19
DB: 3 Gaps: 4

US-10-768-158-1 (1-2419) x US-09-898-570-40 (1-283)
Qy 81 GAGTTCATCGCTGCGGCTGCGCCCTTCTGCGCGGAAGATGGAGGATCGCCAAC 140
Db 1 GluValAsnGlyIleLeuMetSerLysLeuMetSerAspAsnTrpAspLysIleTrpAsn 20
Qy 141 TTCCTCGGTGGCGGCGGAGCGACGCTGTGTGATCGTCACTTACCTACCCCAAGTCCGGCAGCAGCTTG 200
Db 21 PheGlnAlaLysProAspAspLeuLeuAlaThrTyrAlaLysAlaGlyThrThr 40
Qy 201 CTCGAGAGTGTCTACTTGTGTGAGCCAGGCGGTGACCCCGATGAGATGGCTTGATG 260
Db 41 ThrGlnGluIleValAspMetIleGlnAsnAspGlyAspValGlnLysCysGlnArgAla 60
Qy 261 AACATCGACGAGCAGCTCCCGGTCTGGAGTAC-----CCACAGCGC-----GGC 305
Db 61 AsnThrTyrAspArgHisProPheIleGluTrpThrLeuProSerProLeuAsnSerGly 80
```



```
Db      261 SerGluAspPheAspGluAspTyrArgArgLysMetAlaGlySerAsnIleThrPhe 279
RESULT 35
; Sequence 8, Application US/10199330
; Publication No. US20020182681A1
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL00669DIV-1
; CURRENT APPLICATION NUMBER: US/10/199,330
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-330-8
Alignment Scores:
Pred. No.: 4.9e-26 Length: 283
Score: 419.50 Matches: 91
Percent Similarity: 53.0% Conservative: 57
Best Local Similarity: 32.6% Mismatches: 112
Query Match: 9.5% Indels: 19
DB: 4 Gaps: 4
US-10-768-158-1 (1-2419) x US-10-199-330-8 (1-283)
QY      81 GAGTTCCATGCGTGGCGTCCGCCCTTCTGCGCGGAAGATGGAGGAGATCGCCAAC 140
Db      1 GluValAsnGlyIleLeuMetSerLysLeuMetSerAspAsnTrpAspLysIleTrpAsn 20
QY     141 TTCCCGGTGCGGCCAGCGACGTGTGGATCGTACCTACCCCAAGTCCGGCACCAGCTTG 200
Db     21 PheGlnAlaLysProAspAspLeuIleAlaThrTyrAlaLysAlaGlyThrThrTrp 40
QY     201 CTGACGAGGTGGTCTACTTGGTGAGCCAGGCGGTGACCCCGATGAGATCGGCTTGATG 260
Db     41 ThrGlnGluIleValAspMetIleGlnAsnAspGlyAspValGlnLysCysGlnArgAla 60
QY     261 AACATCGACGACGACGTCCTCGGTCTGAGTAC-----CCACAGCG-----GGC 305
Db     61 AsnThrTyrAspArgHisProPheIleGluTrpThrLeuProSerProLeuAsnSerGly 80
QY     306 CTGCACATCATCAAGGAACCTGACCTCTCCCGCCTCATCAAGAGCCACTGCCCTACCGC 365
Db     81 LeuAspLeuAlaAsnLysMetProSerProArgThrLeuLysThrHisLeuProValHis 100
QY     366 TTCTGCGCCTCTGACCTCCACAATGGAGACTCCAAAGTTCATATATGGTTCGCAACCCC 425
Db    101 MetLeuProSerPheTrpLysGluAsnSerLysIleIleTyrValAlaArgAsnAla 120
QY     426 AAGGATCTGGTGGTCTTATATATCAGTTCCACCGCTCTCTCGGACCATGAGTACCGA 485
Db    121 LysAspCysLeuValSerTyrTyrPheSerArgMetAsnLysMetLeuProAspPro 140
QY     486 GGCACCTTTCAGAAATCTCGCGGAGTTTATGAATGATAGCTGGGCTACGGCTCTCGG 545
Db    141 GlyThrLeuGlyGluTyrIleGlnPheLysAlaGlyLysValLeuTrpGlySerTrp 160
QY     546 TTGAGCAGCGTCAGGAGTTCTGGGAGCAGCCGATGGAAGTCTGCTTTTCTCAAG 605
```

```
Db      161 TyrAspHisValLysGlyTrpTrpAspValLysAspGlnHisArgIleLeuTrpPhe 180
QY     606 TATGAAGACATGTCATCGGACCTGCTGACGATGGTGGAGCAGCTGGCCAGATTCTCTGGG 665
Db     181 TyrGluAspMetLysGluAspProLysArgGluIleLysLysIleAlaLysPheLeuGlu 200
QY     666 GTGCTCTGTGACAAGGCCAGCTGGAAGCCCTGACGAGCAGCTGCCACACAGCTGGTG--- 722
Db     201 LysAspIleSerGluGluValLeuAsnLysIleTyrHisThrSerPheAspValMet 220
QY     723 -----GACCACTGCTCAACGCTGAGGCCCTGCGC----- 752
Db     221 LysGluAsnProMetAlaAsnTyrThrThrLeuProSerSerIleMetAspHisSerIle 240
QY     753 -----GTGGCGCGGGAAGAGTGGGTGGGAAGGACATCTTCACCGTCTCCATG 803
Db     241 SerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAlaGln 260
QY     804 AATGAGAAGTTTGACTTGGTGTTATATAACAGAGATGGAAAGTGTGACCTCACGTTT 860
Db     261 SerGluAspPheAspGluAspTyrArgArgLysMetAlaGlySerAsnIleThrPhe 279
RESULT 36
US-10-199-330-9
; Sequence 9, Application US/10199330
; Publication No. US20020182681A1
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL00669DIV-1
; CURRENT APPLICATION NUMBER: US/10/199,330
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-330-9
Alignment Scores:
Pred. No.: 4.9e-26 Length: 283
Score: 419.50 Matches: 91
Percent Similarity: 53.0% Conservative: 57
Best Local Similarity: 32.6% Mismatches: 112
Query Match: 9.5% Indels: 19
DB: 4 Gaps: 4
US-10-768-158-1 (1-2419) x US-10-199-330-9 (1-283)
QY      81 GAGTTCCATGCGTGGCGTCCGCCCTTCTGCGCGGAAGATGGAGGAGATCGCCAAC 140
Db      1 GluValAsnGlyIleLeuMetSerLysLeuMetSerAspAsnTrpAspLysIleTrpAsn 20
QY     141 TTCCCGGTGCGGCCAGCGACGTGTGGATCGTACCTACCCCAAGTCCCGCACCAGCTTG 200
Db     21 PheGlnAlaLysProAspAspLeuIleAlaThrTyrAlaLysAlaGlyThrThrTrp 40
QY     201 CTGACGAGGTGGTCTACTTGGTGAGCCAGGCGGTGACCCCGATGAGATCGGCTTGATG 260
Db     41 ThrGlnGluIleValAspMetIleGlnAsnAspGlyAspValGlnLysCysGlnArgAla 60
```

```
QY 261 AACATCAGCAGCAGCTCCGGTCTCGAGTAC-----CCACAGCCG-----GGC 305
Db 61 AsnThrTyrAspArgHisProPheIleGluThrThrLeuProSerProLeuAsnSerGly 80
QY 306 CTGGACATCATCAAGGAAGTCACTCCCGCCTCATCAAGAGCCACCTGCCCTACCGC 365
Db 81 LeuAspLeuAlaAsnLysMetProSerProArgThrLeuLysThrHisLeuProValHis 100
QY 366 TTTCTGCCCTCTGACCTCCCAATGGAGACTCCCAAGTCACTATATATGGCTGGCAACCC 425
Db 101 MetLeuProProSerPheThrLysGluValLeuLysValLeuThrValAlaArgAsnAla 120
QY 426 AAGGATCTGGTGGTCTTATTATCAGTTCCACCGCTCTCTCGCGACCATGAGCTACCGA 485
Db 121 LysAspCysLeuValSerTyrTyrPheSerArgMetAsnLysMetLeuProAspPro 140
QY 486 GGCACCTTTCAAGAAATCTCGCGGAGTATTATGAATGATGAAGTGGGCTACGGCTCTCGG 545
Db 141 GlyThrLeuGlyGluTyrIleGluGlnPheLysAlaGlyLysValLeuThrGlySerTrp 160
QY 546 TTTGAGCAGCTGCAGGAGTCTCGGAGCAGCCGATGAGCTCAAGCTGCTTTTCTCAAG 605
Db 161 TyrAspHisValLysGlyThrAspValLysAspGlnHisArgIleLeuTyrLeuPhe 180
QY 606 TATGAAGACATGCATCGGACCTGGTGACGATGTGGAGCAGCTGGCCAGATTCTCGGG 665
Db 181 TyrGluAspMetLysGluAspProLysArgGluLeuLysIleAlaLysPheLeuGlu 200
QY 666 GTGTCTCTGTCAAGCCAGCTGGAAGCCCTGAGGAGCAGCTGCCACCTGGTG--- 722
Db 201 LysAspIleSerGluGluValLeuAsnLysIleIleTyrHisThrSerPheAspValMet 220
QY 723 -----GACCAGTGTGCAAGCTGAGCCCTGCCC----- 752
Db 221 LysGluAsnProMetAlaAsnTyrThrThrLeuProSerSerIleMetAspHisSerIle 240
QY 753 -----GTGGCCGGGAGAGTTCGGCTGTGGAGGACATCTTCACCGTCTCCCATG 803
Db 241 SerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAlaGln 260
```

RESULT 37

```
US-10-199-334-8
; Sequence 8, Application US/10199334
; Publication No. US20030008354A1
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-3
; CURRENT APPLICATION NUMBER: US/10/199,334
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-334-8
Alignment Scores:
```

```
Pred. No.: 4.9e-26 Length: 283
Score: 419.50 Matches: 91
Percent Similarity: 53.0% Conservative: 57
Best Local Similarity: 53.6% Mismatches: 112
Query Match: 9.5% Indels: 19
DB: 4 Gaps: 4

US-10-768-158-1 (1-2419) x US-10-199-334-8 (1-283)
QY 81 GAGTTCATCGGTGGCTGGCGCCCTTCTGCGCGGGAAGATGGAGGATCGCCCAAC 140
Db 1 GluValAsnGlyIleLeuMetSerLysLeuMetSerAspAsnTrpAspLysIleTrpAsn 20
QY 141 TTCCCGGTGGCGCCAGCAGCTGGTGGATCGTCACTTACCCCAAGTCCCGCACCAGCTTG 200
Db 21 PheGlnAlaLysProAspAspLeuLeuIleAlaThrTyrAlaLysAlaGlyThrTrp 40
QY 201 CTGCAAGAGTGGTCTACTTGGTGACGCCAGGGGCTGACCCCGATGAGATCGGCTTGATG 260
Db 41 ThrGlnGluIleValAspMetIleGlnAsnAspGlyAspValGlnLysCysGlnArgAla 60
QY 261 AACATCAGCAGCAGCTCCCGCTCTCGAGTAC-----CCACAGCCG-----GGC 305
Db 61 AsnThrTyrAspArgHisProPheIleGluThrThrLeuProSerProLeuAsnSerGly 80
QY 306 CTGACACATCATCAAGGAAGTCACTCTCCCGCCTCATCAAGAGCAGCCACCTGCCCTACCGC 365
Db 81 LeuAspLeuAlaAsnLysMetProSerProArgThrLeuLysThrHisLeuProValHis 100
QY 366 TTTCTGCCCTCTGACCTCCCAATGGAGACTCCAAGGTCACTATATATGGCTGGCAACCC 425
Db 101 MetLeuProProSerPheThrLysGluAsnSerLysIleIleTyrValAlaArgAsnAla 120
QY 426 AAGATCTGTGGTGTCTTATTATCAGTTCACCGCTCTCTCGGACCATGAGCTACCGA 485
Db 121 LysAspCysLeuValSerTyrTyrPheSerArgMetAsnLysMetLeuProAspPro 140
QY 486 GGCACCTTTCAAGAAATCTCCCGGAGTATTATGAATGATGAAGTGGCTGCGCTCTCGG 545
Db 141 GlyThrLeuGlyGluTyrIleGluGlnPheLysAlaGlyLysValLeuThrGlySerTrp 160
QY 546 TTTGAGCAGCTGCAGGAGTTCGGGAGCAGCCGATGAGCTGCAACGTGCTTTTCTCAAG 605
Db 161 TyrAspHisValLysGlyThrAspValLysAspGlnHisArgIleLeuTyrLeuPhe 180
QY 606 TATGAAGACATGCATCGGACCTGGTGACGATGTGGAGCAGCTGGCCAGATTCTCGGG 665
Db 181 TyrGluAspMetLysGluAspProLysArgGluLeuLysIleAlaLysPheLeuGlu 200
QY 666 GTGTCTCTGTGACAGGCCAGCTGGAAGCCCTGAGGAGCAGCTGCCACCTGGTG--- 722
Db 201 LysAspIleSerGluGluValLeuAsnLysIleIleTyrHisThrSerPheAspValMet 220
QY 723 -----GACCAGTGTGCAAGCTGAGCCCTGCCC----- 752
Db 221 LysGluAsnProMetAlaAsnTyrThrThrLeuProSerSerIleMetAspHisSerIle 240
QY 753 -----GTGGCCGGGAGAGTTCGGCTGTGGAGGACATCTTCACCGTCTCCCATG 803
Db 241 SerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAlaGln 260
QY 804 AATGAGAAAGTTTGACTTGGTGTATATAACAGAAAGATGGAAAAGTGTGACCTCACGTTT 860
Db 261 SerGluAspPheAspGluAspTyrArgArgLysMetAlaGlySerAsnIleThrPhe 279

RESULT 38
US-10-199-334-9
; Sequence 9, Application US/10199334
; Publication No. US20030008354A1
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
```

; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF

; FILE REFERENCE: CL00069DIV-3
 ; CURRENT APPLICATION NUMBER: US/10/199,334
 ; CURRENT FILING DATE: 2002-07-22
 ; PRIOR APPLICATION NUMBER: 09/609,816
 ; PRIOR FILING DATE: 2001-03-27
 ; PRIOR APPLICATION NUMBER: 60/192,408
 ; PRIOR FILING DATE: 2000-03-27
 ; PRIOR APPLICATION NUMBER: 60/212,725
 ; PRIOR FILING DATE: 2000-06-20
 ; PRIOR APPLICATION NUMBER: 09/609,816
 ; PRIOR FILING DATE: 2000-07-03
 ; NUMBER OF SEQ ID NOS: 23
 ; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 9

; LENGTH: 283

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-199-334-9

Alignment Scores:

Pred. No.: 4.9e-26 Length: 283
 Score: 419.50 Matches: 91
 Percent Similarity: 53.0% Conservative: 57
 Best Local Similarity: 32.6% Mismatches: 112
 Query Match: 9.5% Indels: 19
 DB: 4 Gaps: 4

US-10-768-158-1 (1-2419) x US-10-199-334-9 (1-283)

QY 81 GAGTTCATGCGTGGCTCGGCCCTTCTGCGCGGGAAGATGGAGGATCGCCAAC 140
 DB 1 GluValAsnGlyIleLeuMetSerLysLeuMetSerAspAsnTrpAspLysIleTrpAsn 20
 QY 141 TTCCCGGTGCGCCAGCAGCTGTGGATCGTCACTACCCCAAGTCGGCCAGCAGCTTG 200
 DB 21 PheGlnAlaLysProAspAspLeuIleAlaThrTyraLysAlaGlyThrTrp 40
 QY 201 CTGCAGGAGGTGTACTTGGTGAGCCAGGCGCTGACCCCGATGAGATCGGTTGATG 260
 DB 41 ThrGlnGluIleValAspMetIleGlnAsnAspGlyAspValGlnLysCysGlnArgAla 60
 QY 261 AACATCGACGAGCTCCCGGTCTGGAGTAC-----CCACAGCCG-----GGC 305
 DB 61 AsnThrTyraAspArgHisProPheIleGluTrpThrLeuProSerProLeuAsnSerGly 80
 QY 306 CTGCACATCATCAAGGAAGTCACTCTCCCGCGCTCATCAAGAGCCACCTCCCTACCGC 365
 DB 81 LeuAspLeuAlaAsnLysMetProSerProArgThrLeuLysThrHisLeuProValHis 100
 QY 366 TTCTGCGCCTCTGACCTCCCAATGGAGATCCAAGGTCACTATATGGCTCGCAACCCC 425
 DB 101 MetLeuProProSerPheTrpLysGluAsnSerLysIleIleTyraValAlaArgAsnAla 120
 QY 426 AAGATCTGGTGTCTTATATCAGTTCCACCGCTCTCTCGGACCATGAGTACCGA 485
 DB 121 LysAspCysLeuValSerTyraTyraPheSerArgMetAsnLysMetLeuProAspPro 140
 QY 486 GGCACCTTTCAAGAAATCTCGCGGAGTTTATGATGATAGCTGGGCTACGCTCTCTGG 545
 DB 141 GlyThrLeuGlyGluTyraIleGluGlnPheLysAlaGlyLysValLeuTrpGlySerTrp 160
 QY 546 TTGAGCAGCTGCAGGAGTCTCGGAGCAGCCGATGAGTCAAGCTGCTTTTCTCAAG 605
 DB 161 TyrAspHisValLysGlyTrpTrpAspValLysAspGlnHisArgIleLeuTyraPhe 180
 QY 606 TATGAGACATGCATCGGAGCTGGTACGATGTGGAGCAGCTGGCGAGATTCCTGGGG 665
 DB 181 TyrGluAspMetLysGluAspProLysArgGluIleLysLysIleAlaLysPheLeuGlu 200
 QY 666 GTGTCCTGTGACAGGCCAGCTGGAAGCCCTGACGGAGCATGCCACCAGCTGGTG--- 722
 DB 201 LysAspIleSerGluGluValLeuAsnLysIleIleTyraHisThrSerPheAspValMet 220

QY 723 -----GACCAGTGTCTCAACGCTGAGGCCCTGCC----- 752
 DB 221 LysGluAsnProMetAlaAsnTyThrThrLeuProSerSerIleMetAspHisSerIle 240
 QY 753 -----GTGGCCCGGGAAGAGTGGCGCTGGGAAGGACATCTTCACCGTCTCCATG 803
 DB 241 SerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyraPheThrValAlaGln 260
 QY 804 AATGAGAAAGTTTGACTTGGTGATATAAACAAGATGGAAAGTGTGACCTCACGTTT 860
 DB 261 SerGluAspPheAspGluAspTyraArgLysMetAlaGlySerAsnIleThrPhe 279

RESULT 39

US-10-199-329-8

; Sequence 8, Application US/10199329

; Publication No. US20030166189A1

; GENERAL INFORMATION:

; APPLICANT: WOODAGE, Trevor et al.

; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING

; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN

; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF

; FILE REFERENCE: CL00069DIV-2

; CURRENT APPLICATION NUMBER: US/10/199,329

; CURRENT FILING DATE: 2002-07-22

; PRIOR APPLICATION NUMBER: 09/609,816

; PRIOR FILING DATE: 2001-03-27

; PRIOR APPLICATION NUMBER: 60/192,408

; PRIOR FILING DATE: 2000-03-27

; PRIOR APPLICATION NUMBER: 60/212,725

; PRIOR FILING DATE: 2000-06-20

; PRIOR APPLICATION NUMBER: 09/609,816

; PRIOR FILING DATE: 2000-07-03

; NUMBER OF SEQ ID NOS: 23

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 8

; LENGTH: 283

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-199-329-8

Alignment Scores:

Pred. No.: 4.9e-26 Length: 283
 Score: 419.50 Matches: 91
 Percent Similarity: 53.0% Conservative: 57
 Best Local Similarity: 32.6% Mismatches: 112
 Query Match: 9.5% Indels: 19
 DB: 4 Gaps: 4

US-10-768-158-1 (1-2419) x US-10-199-329-8 (1-283)

QY 81 GAGTTCATGCGTGGCTCGGCCCTTCTGCGCGGGAAGATGGAGGATCGCCAAC 140
 DB 1 GluValAsnGlyIleLeuMetSerLysLeuMetSerAspAsnTrpAspLysIleTrpAsn 20
 QY 141 TTCCCGGTGCGCCAGCAGCTGTGGATCGTCACTACCCCAAGTCGGCCAGCAGCTTG 200
 DB 21 PheGlnAlaLysProAspAspLeuIleAlaThrTyraLysAlaGlyThrTrp 40
 QY 201 CTGCAGGAGGTGTCTACTTGGTGAGCCAGGCGCTGACCCCGATGAGATCGGTTGATG 260
 DB 41 ThrGlnGluIleValAspMetIleGlnAsnAspGlyAspValGlnLysCysGlnArgAla 60
 QY 261 AACATCGACGAGCTCCCGGTCTGGAGTAC-----CCACAGCCG-----GGC 305
 DB 61 AsnThrTyraAspArgHisProPheIleGluTrpThrLeuProSerProLeuAsnSerGly 80
 QY 306 CTGCACATCATCAAGGAAGTCACTCTCCCGCGCTCATCAAGAGCCACCTCCCTACCGC 365
 DB 81 LeuAspLeuAlaAsnLysMetProSerProArgThrLeuLysThrHisLeuProValHis 100
 QY 366 TTCTGCGCCTCTGACCTCCCAATGGAGATCCAAGGTCACTATATGGCTCGCAACCCC 425


```

Db      41 ThrGlnGluIleValAspMetIleGlnAsnAspGlyAspValGlnLysCysGlnArgAla 60
Qy      261 AACATCGACGAGCAGCTCCCGGTCTCTGGAGTACCCACAGCCG-----GGC 305
Db      61 AsnThrTyrAspArgHisPheIleGluTrpThrLeuProProLeuAsnSerGly 80
Qy      306 CTGGACATCATCAAGGAAGTCACTCCCGCTCATCAAGACCCACCTGCGCTACCGC 365
Db      81 LeuAspLeuAlaAenLysMetProSerProArgThrLeuLysThrHisLeuProValGln 100
Qy      366 TTTCTGCGCTCTGACCTCCACAATGGAGACCTCAAGGTCACTATATGGTTCGCAACCC 425
Db      101 MetLeuProProSerPheTrpLysGluAsnSerGlnIleIleTyrValAlaArgAsnAla 120
Qy      426 AAGGATCTGGTGTCTTATTATCATGTTCCACCGCTCTCTCGGACCATGAGTACCGA 485
Db      121 LysAspCysLeuValSerTyrTyrPheSerArgMetAenLysMetLeuProAspPro 140
Qy      486 GGCACCTTTCAGAAATCTGCGGAGGTATTATGATATGATAGCTGGGCTACGGCTCTGG 545
Db      141 GlyThrLeuGlyGluTyrIleGluThrPheLysAlaGlyLysValLeuTrpGlySerTrp 160
Qy      546 TTTGAGCAGCTGCGAGGAGTCTGGGAGCACCGCATGGACTCGAAGCTGCTTTTCTCAAG 605
Db      161 TyrAspHisValLysGlyTrpTrpAspValLysAspLysHisArgIleLeuTyrLeuPhe 180
Qy      606 TATGAAGACATGATCGGAGCTCTGGTACGATGGTGGAGCAGCTGGCCAGATCTCTGGGG 665
Db      181 TyrGluAspMetLysGluAspProLysArgGluIleLysValLysPheLeuGlu 200
Qy      666 GTGCTCTGTGACAGGAGTGTGTTATTAACAGAGATGGGAAGTGTGACCTCACGTTT 860
Db      201 LysAspLeuSerGluGluValLeuAsnLysIleIleHisHisThrSerPheAspValMet 220
Qy      723 -----GACCATGCTGCAAGCTGAGCGCTGCGC----- 752
Db      221 LysGlnAsnProMetAlaAsnTyrThrLeuProSerSerIleMetAspHisSerIle 240
Qy      753 -----GTGGCCGGGGAAGTGTGGGTGGAGGAGCATCTTCCACGCTCTCCATG 803
Db      241 SerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAlaGln 260

```

RESULT 46

```

US-10-199-329-10
; Sequence 10, Application US/10199329
; Publication No. US20030166189A1
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-2
; CURRENT APPLICATION NUMBER: US/10/199,329
; CURRENT FILING DATE: 2002-07-22
; PRIOR FILING DATE: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR FILING DATE: 2000-03-27
; PRIOR FILING DATE: 2000-06-20
; PRIOR FILING DATE: 2000-06-20
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Homo sapiens

```

US-10-199-329-10

Alignment Scores:

```

Pred. No.:      4,13e-25      Length:      283
Score:          408.50      Matches:      88
Percent Similarity: 52.0%      Conservative: 57
Best Local Similarity: 31.5%      Mismatches: 115
Query Match:     9.3%      Indels:      19
DB:              4      Gaps:      3

```

US-10-768-158-1 (1-2419) x US-10-199-329-10 (1-283)

```

Qy      81 GAGTTCATCGCTGCGGTCTCGCCCTTCTGCGGGGAAGATGGAGGATCGCCAAC 140
Db      1 GluValAsnGlyIleLeuMetSerLysMetSerGluAsnTrpAspLysIleTrpAsn 20
Qy      141 TTCCCGGTGGCGGCCAGCGAGCTGTGGATCGTCACTTACCCCAAGTCCGGCACCAGCTTG 200
Db      21 PheGlnAlaLysProAspAspLeuLeuIleAlaThrTyrAlaLysAlaGlyThrThrTrp 40
Qy      201 CTGACAGAGTGTCTACTTGTGTGACGAGCGGCTGACCCCGATGATGATCGGCTTGATG 260
Db      41 ThrGlnGluIleValAspMetIleGlnAsnAspGlyAspValGlnLysCysGlnArgAla 60
Qy      261 AACATCGACGAGCAGCTCCCGGTCTCTGGAGTACCCACAGCCG-----GGC 305
Db      61 AsnThrTyrAspArgHisPropheIleGluTrpThrLeuProProLeuAsnSerGly 80
Qy      306 CTGGACATCATCAAGGAAGTCACTCTCCCGCTCATCAAGACCCACCTGCGCTTCCATCCGC 365
Db      81 LeuAspLeuAlaAenLysMetProSerProArgThrLeuLysThrHisLeuProValGln 100
Qy      366 TTTCTGCGCTCTGACCTCCACAATGAGACTCCAAAGTCACTATATGGTTCGCAACCC 425
Db      101 MetLeuProProSerPheTrpLysGluAsnSerGlnIleIleTyrValAlaArgAsnAla 120
Qy      426 AAGGATCTGTGTGCTTATTATCATGTTCCACCGCTCTCTGGGACCATGAGTACCGA 485
Db      121 LysAspCysLeuValSerTyrTyrPheSerArgMetAenLysMetLeuProAspPro 140
Qy      486 GGCACCTTTCAGAAATCTGCGGAGGTATTATGAATGATAGCTGGGCTCGGCTCTCCATG 545
Db      141 GlyThrLeuGlyGluTyrIleGluThrPheLysAlaGlyLysValLeuTrpGlySerTrp 160
Qy      546 TTTGAGCAGCTGCGAGGAGTCTGGGAGCACCGCATGGACTCGAAGCTGCTTTTCTCAAG 605
Db      161 TyrAspHisValLysGlyTrpTrpAspValLysAspLysHisArgIleLeuTyrLeuPhe 180
Qy      606 TATGAAGACATGATCGGAGCTGCTGACGATGGTGGAGCAGCTGGCCAGATCTCTGGGG 665
Db      181 TyrGluAspMetLysGluAspProLysArgGluIleLysValLysPheLeuGlu 200
Qy      666 GTGCTCTGTGACAGGAGTGTGAAAGCTGGAAGCTGCAAGGACACTGCCACAGCTGGTG--- 722
Db      201 LysAspIleSerGluGluValLeuAsnLysIleIleHisHisThrSerPheAspValMet 220
Qy      723 -----GACCATGCTGCAAGCTGAGCGCTGCGC----- 752
Db      221 LysGlnAsnProMetAlaAsnTyrThrLeuProSerSerIleMetAspHisSerIle 240
Qy      753 -----GTGGCCGGGGAAGTGTGGGTGGAGGAGCATCTTCCACGCTCTCCATG 803
Db      241 SerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAlaGln 260
Qy      804 AATGAGAAGTTTGACTTGTGTATATAACAGAGATGGGAAGTGTGACCTCACGTTT 860
Db      261 SerGluAspPheAspGluAspTyrArgLysLysMetAlaGlySerThrIleThrPhe 279

```

RESULT 47

```

US-11-108-875-10
; Sequence 10, Application US/11108875
; Publication No. US20050196802A1
; GENERAL INFORMATION:

```

```
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-4
; CURRENT APPLICATION NUMBER: US/11/108,875
; CURRENT FILING DATE: 2005-04-19
; PRIOR APPLICATION NUMBER: 10/199,334
; PRIOR FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-108-875-10

Alignment Scores:
Pred. No.: 4.13e-25 Length: 283
Score: 408.50 Matches: 88
Percent Similarity: 52.0% Conservative: 57
Best Local Similarity: 31.5% Mismatches: 115
Query Match: 9.3% Indels: 19
DB: 6 Gaps: 3

US-10-768-158-1 (1-2419) x US-11-108-875-10 (1-283)
QY 81 GAGTTCATCGCGTGGCGCTCGCCCTTCGCGCGGAAGATGGAGAGATCGCCAAC 140
DB 1 GluValAsnGlyIleLeuMetSerLysMetMetSerGluAsnTrpAspLysIleTrpAsn 20
QY 141 TTCCCGTGGCGGCCAGCGAGTGTGGTGTACCTACCCAGTCCCGGACCGAGTTG 200
DB 21 PheGlnAlaLysProAspAspLeuLeuIleAlaThrTyrAlaLysAlaGlyThrThrTrp 40
QY 201 CTGCAGAGGTGTCTACTTGGTGGAGGCGGCGCTGACCCCGATGAGTGGGTTGATG 260
DB 41 ThrGlnIleValAspMetIleGlnAsnAspGlyAspValGlnLysCysGlnArgAla 60
QY 261 AACATCGACGAGCGTCCCGGCTCTCGAGTACCACACAGCG-----GGC 305
DB 61 AsnThrTyrAspArgHisProPheIleGluTrpThrLeuProProLeuAsnSerGly 80
QY 306 CTGGACATCATCAAGAACTGACCTCTCCCGCTCATCAAGAGCCACCTGCCCTACCGC 365
DB 81 LeuAspLeuAlaAsnLysMetProSerProArgThrLeuLysThrHisLeuProValGln 100
QY 366 TTCTGCGCTCTGACCTCCACAAATGGAGATCCAGGTCACTATATGGTTCGCAACCCC 425
DB 101 MetLeuProProSerPheTrpLysGluAsnSerGlnIleIleTrpValAlaArgAsnAla 120
QY 426 AAGATCTGGTGGTGTCTATTATCAGTTCACCGCTCTCTGCGGACCATGAGTACCGA 485
DB 121 LysAspCysLeuValSerTyrTyrPheSerArgMetAsnLysMetLeuProAspPro 140
QY 486 GGCACTTTCAGAAATTCGCGGAGGTTTATGAATGATAGTGGCTACGGCTCCTCGG 545
DB 141 GlyThrLeuGlyGluTyrIleGluThrPheLysAlaGlyLysValLeuTrpGlySerTrp 160
QY 546 TTTGAGACGTCGAGGAGTCTTGGGACCGCATGAGTCAACAGTGCCTTTTCTCAAG 605
DB 161 TyrAspHisValLysGlyTrpTrpAspValLysAspLysHisArgIleLeuTyrLeuPhe 180
QY 606 TATGAACACATGTCGCGGACCTGGTACGATGGTGGAGAGCTGGCAGATTCTCTGGG 665
DB 181 TyrGluAspMetLysGluAspProLysArgGluIleLysIleValLysPheLeuGlu 200
```

```
QY 666 GTGTCCTGTGACAAAGCCCGAGCTGGAGAGCCCTGACGGAGCAGCTGCCACAGCTGGTG--- 722
DB 201 LysAspLysSerGluGluValLeuAsnLysIleHisThrSerPheAspValMet 220
QY 723 -----GACCAGTGTGCAACGCTGAGGCGCCCTGCCC----- 752
DB 221 LysGlnAsnProMetAlaAsnTyrThrLeuProSerSerIleMetAspHisSerIle 240
QY 753 -----GTGGCGCGGGAAGAGTTGGCTGTGAAGAGACATCTTTCACCGTCTCCATG 803
DB 241 SerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAlaGln 260
QY 804 AATGAGAAGTTTCACTTGGTGTATAAACAGAAAGTGGAAAGTGTGACCTCACGTTT 860
DB 261 SerGluAspPheAspGluAspTyrArgLysLysMetAlaGlySerThrIleThrPhe 279

RESULT 48
US-10-072-012-584
; Sequence 584, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Beba
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 584
; LENGTH: 304
```



```
Db 181 TrpPheAspHisValLysGlyTrpTrpAlaAlaLysAspThrHisArgIleLeuTyrLeu 200
Qy 603 AAGTATGAAGACATCGCATCGGAGCCTGGTACGATGGTGGAGCAGCTGGCCAGATTCTCTG 662
Db 201 PheTyrGluAspIleLysLysAsnProLysHisGluIleHisLysValLeuGluPheLeu 220
Qy 663 GGGGTGCTCTGTGACAAAGCCAGCTGGAAGCCCTGACGGAGCACTGCCACACAGCTGGTG 722
Db 221 GluLysThrLeuSerGlyAspValIleAsnLysIleValHisHisThrSerPheAspVal 240
Qy 723 -----GACAGTGTGCAAGCTGAGGCTGAGCCCTGCCCC----- 752
Db 241 MetLysAspAsnProMetAlaAsnHisThrAlaValProAlaHisIlePheAsnHisSer 260
Qy 753 -----GTGGCCGGGGAAGATTGGGCTGTGGAAGCAGCATCTTCACCGTCTCC 800
Db 261 IleSerLysPheMetArgLysGlyMetProGlyAspTrpLysAsnHisPheThrValAla 280
Qy 801 ATGAATGGAAGTTGACTTGGTGTATAAACAGAGATGGGAAGTGTGACCTCACGTTT 860
Db 281 MetAsnGluAsnPheAspLysHisTyrGluLysLysMetAlaGlySerThrLeuAsnPhe 300
RESULT 50
US-10-199-334-6
; Sequence 6, Application US/10199334
; Publication No. US20030008354A1
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-3
; CURRENT APPLICATION NUMBER: US/10/199,334
; CURRENT FILING DATE: 2002-07-22
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-334-6
Alignment Scores:
Pred. No.: 5,67e-25 Length: 304
Score: 407.00 Matches: 94
Percent Similarity: 51.7% Conservative: 61
Best Local Similarity: 31.3% Mismatches: 125
Query Match: 9.2% Indels: 20
DB: 4 Gaps: 5
US-10-768-158-1 (1-2419) x US-10-199-334-6 (1-304)
Qy 21 ATGGCGGAGACGAGCGGAGACCCCGAGACCCCGGGGAGTTCGAG---AGCAAGTAC 77
Db 1 MetAlaLysIleGluLysAsnAlaProThrMetGluLysLysProGluLeuPheAsnIle 20
Qy 78 TTCGAGTTCATCGCTGGCTGGCTGCCGCTTCTGCCGCGGGAAGATGAGAGATCGCC 137
Db 21 MetGluValAspGlyValProThrLeuLeuSerLysGluTyrTrpGluLysValCys 40
Qy 138 AACTTCCCGGTGGCGGCGAGCGAGCTGTGGATCGTCACCTACCCCAAGTCCGGACACGAGC 197
Db 41 AsnPheGlnAlaLysProAspLysIleLeuAlaThrTyrProLysSerGlyThrThr 60
Qy 198 TTGTCGAGGAGGTGGTCTACTTGGTGTAGCCAGGCGCTGACCCCGATGAGATCGGCTTG 257
```

```
Db 61 TrpMethHisGluIleLeuAspMetIleLeuAsnAspGlyAspValGluLysCysLysArg 80
Qy 258 ATGAACATCGACGAGCAGCTCCCGGTCTCTGAG-----TACCCA-----CAGCCG 302
Db 81 AlaGlnThrLeuAspArgHisAlaPheLeuGluLeuLysPheProHisLysGluLysPro 100
Qy 303 GGCCTGACATCATCAAGGAATGACCTCTCCCGCTCATCAAGAGCCACTGCCCTAC 362
Db 101 AspLeuGluPheValLeuGluMetSerSerProGlnLeuIleLysThrHisLeuProSer 120
Qy 363 CGCTTTCTGCCCTCTGACCTCCCAATGGAGACTCCCAAGGTCACTCATATATGGCTCGCAAC 422
Db 121 HisLeuIleProProSerIleTrpLysGluAsnCysLysIleValTyrValAlaArgAsn 140
Qy 423 CCAAGATATCTGGTGGTGTCTTATTATCAGTTCCACCGCTCTCTCGCGACCATGAGTAC 482
Db 141 ProLysAspCysLeuValSerTyrHisPheHisArgMetAlaSerPheMetProAsp 160
Qy 483 CGAGGCACCTTTCAAGAAATTCGCCGAGGTTTATGAATGATAAGCTGGGTACGGCTCC 542
Db 161 ProGlnAsnLeuGluPheTyrGluLysPheMetSerGlyLysValValGlyGlySer 180
Qy 543 TGGTTTCAGCACGTGACGAGGTTCTGGAGCACCGCATGGACTCGAAGCTGCTTTTCTC 602
Db 181 TrpPheAspHisValLysGlyTrpTrpAlaAlaLysAspThrHisArgIleLeuTyrLeu 200
Qy 603 AAGTATCAAGACATGTCATCGGACCTGTCAGCATGGTGGAGCAGCTGGCCAGATTCCTG 662
Db 201 PheTyrGluAspIleLysLysAsnProLysHisGluIleHisLysValLeuGluPheLeu 220
Qy 663 GGGGTGCTCTGTGACAAAGCCAGCTGGAAGCCCTGACGGAGCACTGCCACACAGCTGGT 722
Db 221 GluLysThrLeuSerGlyAspValIleAsnLysIleValHisHisThrSerPheAspVal 240
Qy 723 -----GACCAGTCTCTGCAAGCTGAGCCCTGCCCC----- 752
Db 241 MetLysAspAsnProMetAlaAsnHisThrAlaValProAlaHisIlePheAsnHisSer 260
Qy 753 -----GTGGCCCGGGAAGATTGGGCTGTGGAAGGACATCTTCACCGTCTCC 800
Db 261 IleSerLysPheMetArgLysGlyMetProGlyAspTrpLysAsnHisPheThrValAla 280
Qy 801 ATGAATGGAAGTTGACTTGGTGTATAAACAGAGATGGGAAGTGTGACCTCACGTTT 860
Db 281 MetAsnGluAsnPheAspLysHisTyrGluLysLysMetAlaGlySerThrLeuAsnPhe 300
RESULT 51
US-10-199-329-6
; Sequence 6, Application US/10199329
; Publication No. US20030166189A1
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-2
; CURRENT APPLICATION NUMBER: US/10/199,329
; CURRENT FILING DATE: 2002-07-22
; PRIOR FILING DATE: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
```

US-10-199-329-6

Alignment Scores:		
Pred. No.:	5,678-25	304
Score:	407.00	94
Percent Similarity:	51.7%	Conservative: 61
Best Local Similarity:	31.3%	Mismatches: 125
Query Match:	9.2%	Indels: 20
DB:	4	Gaps: 5

US-10-768-158-1 (1-2419) x US-10-199-329-6 (1-304)

Qy	21	ATGGCGGAGACGAGCGACGACCCCGGAGAGTTTCGAG---AGCAAGTAC	77
Db	1	MetAlaIleGluLeuAsnAlaProThrMetGluLysLysProGluLeuPheAsnIle	20
Qy	78	TTGAGTTCCATTGGCTGCGGCTGCGCCCTTCTCCGCGGGAAGATGGAGAGATCGCC	137
Db	21	MetGluAlaAspGlyValProThrLeuIleLeuSerLysGluTrpTrpGluLysValCys	40
Qy	138	AACTTCCCGGTGCGGCCCGACGACGTGTGGATCGTCACTACCCCAAGTCGCGCACCGC	197
Db	41	AsnPheGlnAlaLysProAspAspLeuIleLeuAlaThrTyProLysSerGlyThrThr	60
Qy	198	TTGCTGCAGGAGTGGTCTACTTGTGTGACCGAGCGGCTGACCCCGATCGATCGGCTTG	257
Db	61	TrpMetHisGluIleLeuAspMetIleLeuAsnAspGlyAspValGluLysCysLysArg	80
Qy	258	ATGAACATCGACGACGACGCTCCCGTCTCTGGAG-----TACCCA-----CAGCGG	302
Db	81	AlaGlnThrLeuAspArgHisAlaPheLeuGluLeuLysPheProHisLysGluLysPro	100
Qy	303	GGCTGGACATCATCAAGGAATGACCTCTCCCGCTCATCAAGAGCCACCTGGCCTTAC	362
Db	101	AspLeuGluPheValLeuGluMetSerProGlnLeuIleLysThrHisLeuProSer	120
Qy	363	CGCTTCTGCCCTCTGCTCCACATGGAGACTCCAAGTGCATCTATATGCTCGCAAC	422
Db	121	HisLeuIleProProSerIleTrpLysGluAsnCysLysIleValTyValAlaArgAsn	140
Qy	423	CCCAAGGATCTGGTGTCTTATTATCAGTTCCACCGCTCTCTCGGACCATGAGCTAC	482
Db	141	ProLysAspCysLeuValSerTyTrpHisPheHisArgMetAlaSerPheMetProAsp	160
Qy	493	CGAGGCACCTTTCAGAAATCTGCCGAGGTTTATGAATGATGAAGCTGGCTACGGCTCC	542
Db	161	ProGlnAsnLeuGluGluPheTyTrpLysPheMetSerGlyLysValValGlyGlySer	180
Qy	543	TGTTTTGAGCAGCTGCAGAGTCTTGGGAGCACCGCATGGACTCGAAACGTGTTTTCTC	602
Db	181	TrpPheAspHisValLysGlyTrpTrpAlaAlaLysAspThrHisArgIleLeuTyLeu	200
Qy	603	AAGTATGAAGACATGCTCGGACCTGTTGACGATGGTGGACAGCTGCCAGATCCCTG	662
Db	201	PheTyGluAspIleLysLysAsnProLysHisGluIleHisLysValLeuGluPheLeu	220
Qy	663	GGGGTGTCCTGTGCAAGGCCCGCTGGAAGCCCTGACGAGCATGCCACGAGTGGTG	722
Db	221	GluLysThrLeuSerGlyAspValIleAsnLysIleValHisHisThrSerPheAspVal	240
Qy	723	-----GACCATGTGTCAACGCTGAGGCCCTGCCCC-----752	
Db	241	MetLysAspAsnProMetAlaAsnHisThrAlaValProAlaHisIlePheAsnHisSer	260
Qy	753	-----GTGCGCGCGGGAAGAGTTGGCTGTGGAAGGACATCTTACCCTGCTCC	800
Db	261	IleSerLysPheMetCargLysGlyMetProGlyAspTrpLysAsnHisPheThrValAla	280
Qy	801	ATGAATGAGAACTTTGACTTGTGTATAAACAAGAAGATGGGAAGAGTGTGACCTCAGTTT	860
Db	281	MetAsnGluAsnPheAspLysHisTyTrpGluLysLysMetAlaGlySerThrLeuAsnPhe	300

RESULT 52

US-11-108-875-6

Sequence 6, Application US/11108875
Publication No. US20050196802A1
GENERAL INFORMATION:
APPLICANT: WOODAGE, Trevor et al.
TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
FILE REFERENCES: CL000669DIV-4
CURRENT APPLICATION NUMBER: US/11/108,875
CURRENT FILING DATE: 2005-04-19
PRIOR APPLICATION NUMBER: 10/199,334
PRIOR FILING DATE: 2002-07-22
PRIOR APPLICATION NUMBER: 09/609,816
PRIOR FILING DATE: 2000-07-03
PRIOR APPLICATION NUMBER: 60/212,725
PRIOR FILING DATE: 2000-06-20
PRIOR APPLICATION NUMBER: 60/192,408
PRIOR FILING DATE: 2000-03-27

Alignment Scores:		
Pred. No.:	5.67e-25	304
Score:	407.00	94
Percent Similarity:	51.7%	Conservative: 61
Best Local Similarity:	31.3%	Mismatches: 125
Query Match:	9.2%	Indels: 20
DB:	6	Gaps: 5

US-10-768-158-1 (1-2419) x US-11-108-875-6 (1-304)

QY	21	ATGCGGAGACGAGCGCGAGACCCCGGAGTTCGAG----	AGCAAGTAC	77
DB	1	MetAlaLysIleGluLysAsnAlaProThrMetGluLysLysProGluLeuPheAsnIle	20	
QY	78	TTCCAGTTCCATGCGTGCGGCTGCCGCCCTTCTCCGCGGGAAGATGGAGGAGATCGCC	137	
DB	21	MetGluValAspGlyValProThrLeuIleLeuSerLysGluTrpGluLysValCys	40	
QY	138	AACTTCCGGTCCGCCCGACGACGTGTGGATCGTCACCTACCCCAAGTCCGGCACCAAGC	197	
DB	41	AsnPheGlnAlaLysProAspAspLeuIleLeuAlaThrTyProLysSerGlyThrThr	60	
QY	198	TTGCTGCAGGAGGTGGTCTACTTGGTGTAGCCAGGCGCTGACCCCGATGAGATCGGCTTG	257	
DB	61	TrpMetHisGluIleLeuAspMetIleLeuAsnAspGlyAspValGluLysCysLysArg	80	
QY	258	ATGAACATCGACGACGACGTCCCGGTCTCTGGAG-----TACCCA-----	CAGCCG	302
DB	81	AlaGlnThrLeuAspArgHisAlaPheLeuGluLeuLysPheProHisLysGluLysPro	100	
QY	303	GGCCTGGACATCATCAAGGAATGACCTCTCCCGCTCATCAAGAGCCACCTGCCCTAC	362	
DB	101	AspLeuGluPheValLeuGluMetSerSerProGlnLeuIleLysThrHisLeuProSer	120	
QY	363	CGCTTTCTGCCTCTGACCTCCACAAATGGAGATCCCAAGGTGCTATATGCTCGCAAC	422	
DB	121	HisLeuIleProProSerIleTrpLysGluAsnCysLysIleValTyValAlaArgAsn	140	
QY	423	CCCAGGATCTGGTGGTCTCTATTATCATGTTCCACCGCTCTCTCGGACCACTAGCTAC	482	
DB	141	ProLysAspCysLeuLysIleSerTyTrpHisPheHisArgMetalSerPheMetProAsp	160	
QY	483	CGAGGCACCTTTCAAGAAATCTGCGGAGGTTTATGAATGATAAGCTGGGCTACGGCTCC	542	
DB	161	ProGlnAsnLeuGluGluPheTytrGluLysPheMetSerGlyLysValValGlyLys	180	


```
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: fastseq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-330-4

Alignment Scores:
Pred. No.: 1.81e-24 Length: 304
Score: 401.00 Matches: 93
Percent Similarity: 51.7% Conservative: 62
Best Local Similarity: 31.0% Mismatches: 125
Query Match: 9.1% Indels: 20
DB: 4 Gaps: 5

US-10-768-158-1 (1-2419) x US-10-199-330-4 (1-304)

QY 21 ATGGCGGAGAGCGAGCGAGACCCCGAGACCCCGGGGAGTTCGAG---AGCAAGTAC 77
Db 1 MetAlaLysIleGluLysAsnAlaProThrMetGluLysLysProGluLeuPheAsnIle 20
QY 78 TTCGAGTTCATCGCGTGGCGTGGCGCCCTCTCTCCCGCGGAGAGATGGAGGAGATCGCC 137
Db 21 MetGluValAspGlyValProThrLeuLeuLeuSerLysGluTrpTrpGluLysValCys 40
QY 138 AACTTCCCGTGGCGGCGGAGCGAGCGTGGATGTCACCTACCCCAAGTCGGGACCGAGC 197
Db 41 AsnPheGlnAlaLysProAspAspLeuLeuAlaThrTyProLysSerGlyThrThr 60
QY 198 TTGCTGCGAGGAGTGGTCTACTTGGTGAGCCAGGCGCTGACCCCGATGAGATCGGCTTG 257
Db 61 TrpMetHisGluIleLeuAspMetIleLeuAsnAspGlyAspValGluLysCysLysArg 80
QY 258 ATGAACATCGACGAGCGAGCTCCCGTCTCTGAG---TACCCA-----CAGCCG 302
Db 81 AlaGlnThrLeuAspArgHisAlaPheLeuGluLysPheProHisLysGluLysPro 100
QY 303 GGCTTGACATCATCAAGAACTGCTCTCCCGCTCATCAAGAGCCACTGCCCTAC 362
Db 101 AspLeuGluPheValLeuGluMetSerSerProGlnLeuIleLysThrHisLeuProSer 120
QY 363 CGCTTTCCTGCGCTGACCTCCCAATCGAGACTCCAGGTCATCATATATGCTCGCAAC 422
Db 121 HisLeuIleProSerIleTrpLysGluAsnCysLysIleValIleValAlaArgAsn 140
QY 423 CCCAAGGATCTGGTGGTGTCTTATTATCAGTTCCACCGCTCTCTCGGACCATGAGTAC 482
Db 141 ProLysAspCysLeuValSerTyTrpHisPheHisArgMetAlaSerPheMetProAsp 160
QY 483 CGAGGACCTTTCAGAAATTCGCGGAGGTTTATGAATGATAGCTGGGCTACGGCTCC 542
Db 161 ProGlnAsnLeuGluGluPheTyTrpLysPheMetSerGlyLysValValGlySer 180
QY 543 TGGTTTGACAGTGCAGAGTTCGCGAGCACCCCATGGATCGAAGCTGCTTTTCTC 602
Db 181 TrpPheAspHisValLysGlyTrpTrpAlaLysAspMetHisArgIleLeuTyTrpLeu 200
QY 603 AAGTATGAAGACATCATCGGAGCGTGGTGGACGATGGTGGAGCAGCTGGCCAGATTCTCTG 662
Db 201 PheTyTrpGluAspIleLysLysAsnProLysHisGluIleHisLysValLeuGluPheLeu 220
QY 663 GGGGTGCTGTGACAAAGCCCGAGCTGGAAGCCCTGACGGAGCATGCGCACCAGCTGGTG 722
Db 221 GluLysThrTrpSerGlyAspValIleAsnLysIleValHisHisThrSerPheAspVal 240
QY 723 -----GACCACTGCTGCAACCTGAGCGCCCTGCCCC----- 752
Db 241 MetLysAspAsnProMetAlaAsnHisThrAlaValProAlaHisIlePheAsnHisSer 260
QY 753 -----GTGGGCGGGGAAAGAGTTGGGCTGTGGGAGGAGCATCTTCCCGTCTCC 800
Db 753 -----GTGGGCGGGGAAAGAGTTGGGCTGTGGGAGGAGCATCTTCCCGTCTCC 800
```

```
Db 261 IleSerLysPheMetArgLysGlyMetProGlyAspTrpLysAsnHisPheThrValAla 280
QY 801 ATGAATCAGAGTTTGACTTGGTGTATATAACAGAGAGATGGGAAAGTGTGACCTCACGTTT 860
Db 281 LeuAsnGluAsnPheAspLysHisTyTrGluLysLysMetAlaGlySerThrLeuAsnPhe 300

RESULT 55
US-10-199-334-4
; Sequence 4, Application US/10199334
; Publication No. US20030008354A1
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-3
; CURRENT APPLICATION NUMBER: US/10/199,334
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: fastseq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-334-4

Alignment Scores:
Pred. No.: 1.81e-24 Length: 304
Score: 401.00 Matches: 93
Percent Similarity: 51.7% Conservative: 62
Best Local Similarity: 31.0% Mismatches: 125
Query Match: 9.1% Indels: 20
DB: 4 Gaps: 5

US-10-768-158-1 (1-2419) x US-10-199-334-4 (1-304)

QY 21 ATGGCGGAGAGCGAGCGAGACCCCGAGACCCCGGGGAGTTCGAG---AGCAAGTAC 77
Db 1 MetAlaLysIleGluLysAsnAlaProThrMetGluLysLysProGluLeuPheAsnIle 20
QY 78 TTCGAGTTCATCGCGTGGCGTGGCGCCCTCTCTCCCGCGGAGAGATGGAGGAGATCGCC 137
Db 21 MetGluValAspGlyValProThrLeuLeuLeuSerLysGluTrpTrpGluLysValCys 40
QY 138 AACTTCCCGTGGCGGCGGAGCGAGCGTGGATGTCACCTACCCCAAGTCGGGACCGAGC 197
Db 41 AsnPheGlnAlaLysProAspAspLeuLeuAlaThrTyProLysSerGlyThrThr 60
QY 198 TTGCTGCGAGGAGTGGTCTACTTGGTGAGCCAGGCGCTGACCCCGATGAGATCGGCTTG 257
Db 61 TrpMetHisGluIleLeuAspMetIleLeuAsnAspGlyAspValGluLysCysLysArg 80
QY 258 ATGAACATCGACGAGCGAGCTCCCGTCTCTGAG---TACCCA-----CAGCCG 302
Db 81 AlaGlnThrLeuAspArgHisAlaPheLeuGluLysPheProHisLysGluLysPro 100
QY 303 GGCTTGACATCATCAAGAACTGCTCTCCCGCTCATCAAGAGCCACTGCCCTAC 362
Db 101 AspLeuGluPheValLeuGluMetSerSerProGlnLeuIleLysThrHisLeuProSer 120
QY 363 CGCTTTCCTGCGCTGACCTCCCAATCGAGACTCCAGGTCATCATATATGCTCGCAAC 422
Db 121 HisLeuIleProSerIleTrpLysGluAsnCysLysIleValIleValAlaArgAsn 140
QY 423 CCCAAGGATCTGGTGGTGTCTTATTATCAGTTCCACCGCTCTCTCGGACCATGAGTAC 482
Db 141 ProLysAspCysLeuValSerTyTrpHisPheHisArgMetAlaSerPheMetProAsp 160
QY 483 CGAGGACCTTTCAGAAATTCGCGGAGGTTTATGAATGATAGCTGGGCTACGGCTCC 542
Db 161 ProGlnAsnLeuGluGluPheTyTrpLysPheMetSerGlyLysValValGlySer 180
QY 543 TGGTTTGACAGTGCAGAGTTCGCGAGCACCCCATGGATCGAAGCTGCTTTTCTC 602
Db 181 TrpPheAspHisValLysGlyTrpTrpAlaLysAspMetHisArgIleLeuTyTrpLeu 200
QY 603 AAGTATGAAGACATCATCGGAGCGTGGTGGACGATGGTGGAGCAGCTGGCCAGATTCTCTG 662
Db 201 PheTyTrpGluAspIleLysLysAsnProLysHisGluIleHisLysValLeuGluPheLeu 220
QY 663 GGGGTGCTGTGACAAAGCCCGAGCTGGAAGCCCTGACGGAGCATGCGCACCAGCTGGTG 722
Db 221 GluLysThrTrpSerGlyAspValIleAsnLysIleValHisHisThrSerPheAspVal 240
QY 723 -----GACCACTGCTGCAACCTGAGCGCCCTGCCCC----- 752
Db 241 MetLysAspAsnProMetAlaAsnHisThrAlaValProAlaHisIlePheAsnHisSer 260
QY 753 -----GTGGGCGGGGAAAGAGTTGGGCTGTGGGAGGAGCATCTTCCCGTCTCC 800
Db 753 -----GTGGGCGGGGAAAGAGTTGGGCTGTGGGAGGAGCATCTTCCCGTCTCC 800
```

```

Db 141 ProlysAspCysLeuValSerTyrHisPheHisArgMetAlaSerPheMetProAsp 160
Qy 483 CGAGGCACTTCAAGAAATTCGCGGAGGTTTATGAATGATAAGCTGGGCTACGGCTCC 542
Db 161 ProGlnAenLeuGluGluPheTyrGluLysPheMetSerGlyLysValValGlySer 180
Qy 543 TGGTTTGAGCAGTCGCGAGGAGTCTCGGAGCACCGCATGGACTGAACGTCTTTTCTC 602
Db 181 TrpPheAspHisValLysGlyTrpTrpAlaAlaLysAspMetHisArgIleLeuTyrLeu 200
Qy 603 AAGTATGAAGACATGCATCGGACCTGCTGACGATGGTGGAGCAGCTGGCAGATTCTCG 662
Db 201 PheTyrGluAspIleLysLysAenProLysHisGluIleHisLysValLeuGluPheLeu 220
Qy 663 GGGGTGCTCTGTGACAAAGCCAGCTGGAAGCCCTGACGGAGCACCTGCCACAGCTGGT 722
Db 221 GluLysThrTrpSerGlyAspValIleAenLysIleValHisHisThrSerPheAspVal 240
Qy 723 -----GACCAGTGTGCAAGCTGCAAGCTGAGCCCTGCGCC----- 752
Db 241 MetLysAspAenProMetAlaAenHisThrAlaValProAlaHisIlePheAenHisSer 260
Qy 753 -----GTGGCCGGGGAAGAGTTGGGCTGTGGAAGGACATCTTCAACGCTCTCC 800
Db 261 IleSerLysPheMetArgLysGlyMetProGlyAspTrpLysAenHisPheThrValala 280
Qy 801 ATGAATGAGAAGTTTGACTTGGTGTATAAACAGAAAGATGGAAAAGTGTGACCTCACGTTT 860
Db 281 LeuAenGluAenPheAspLysHisTyrGluLysLysMetAlaGlySerThrLeuAenPhe 300

```

RESULT 56

```

US-10-199-329-4
; Sequence 4, Application US/10199329
; Publication No. US20030166189A1
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-2
; CURRENT APPLICATION NUMBER: US/10/199,329
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-329-4

```

Alignment Scores:

```

Pred. No.: 1.81e-24 Length: 304
Score: 401.00 Matches: 93
Percent Similarity: 51.7% Conservative: 62
Best Local Similarity: 31.0% Mismatches: 125
Query Match: 9.1% Indels: 20
DB: 4 Gaps: 5

```

US-10-768-158-1 (1-2419) x US-10-199-329-4 (1-304)

```

Qy 21 ATGGCGGAGACGAGCGGCGAGACCCCGGAGAGTTGAG---AGCAAGTAC 77
Db 1 MetAlaLysIleGluLysAenAlaProThrMetGluLysLysProGluLeuPheAenIle 20

```

```

Qy 78 TTGAGTTCCATGCGCTGCGGCTGCCGCCCTTCTGCGCGGGAAGATGGAGGATCGCC 137
Db 21 MetGluValAspGlyValProThrLeuIleLeuSerLysGluTrpTrpGluLysValCys 40
Qy 138 AACTTCCCGGTGCGGCGCCAGCGCTGGATCGTACCTACCCCAAGTCGGCGGACGAGC 197
Db 41 AsnPheGlnAlaLysProAspAspLeuIleuAlaThrTyrProLysSerGlyThrThr 60
Qy 198 TTGCTCGAGGAGTGTCTACTTGTGAGCGAGGCGCTGACCCCGATGAGATCGGTTG 257
Db 61 TrpMethIleGluIleLeuAspMetIleLeuAenAspGlyAspValGluLysCysLysArg 80
Qy 258 ATGAACATCGACGAGCAGCTCCCGCTCTCGAG-----TACCCA-----CAGCCG 302
Db 81 AlaGlnThrLeuAspArgHisAlaPheLeuLeuLysPheProHisLysLysGluLysPro 100
Qy 303 GGCCTGACATCATCAAGGAACGACCTCTCCCGCTCATCAAGAGCCACCTCCCTTAC 362
Db 101 AspLeuGluPheValLeuGluMetSerSerProGlnLeuIleLysThrHisLeuProSer 120
Qy 363 CGCTTCTGCTCTGACCTCCCAATGGAGACTCCAAAGGTCACTATATGGTTCGCAAC 422
Db 121 HisLeuIleProProSerIleTrpLysGluAenCysLysIleValTyrValAlaArgAsn 140
Qy 423 CCAGAGATCTGGTGTGTCTTATTATCAGTTCCACCGCTCTCTCGGAGCAGATGAGTAC 482
Db 141 ProlysAspCysLeuValSerTyrHisPheHisArgMetAlaSerPheMetProAsp 160
Qy 483 CGAGGCACTTTCAGAAATTCGCGGAGGTTTATGAATGATAGCTGGGCTACGGCTCC 542
Db 161 ProGlnAenLeuGluGluPheTyrGluLysPheMetSerGlyLysValValGlySer 180
Qy 543 TGGTTTGAGCAGCTGCGAGGAGTTCTGGGAGCACCGCATGGACTCGAAGCTCTTTTCTC 602
Db 181 TrpPheAspHisValLysGlyTrpTrpAlaAlaLysAspMetHisArgIleLeuTyrLeu 200
Qy 603 AAGTATGAAGACATGCATCGGAGCTCGTGCAGTGTGGAGCAGCTGGCAGATTCCTG 662
Db 201 PheTyrGluAspIleLysLysAenProLysHisGluIleHisLysValLeuGluPheLeu 220
Qy 663 GGGGTGCTGTGACAAAGCCAGCTGGAAGCCCTGACGGAGCAGCTGCCACGAGTGTG 722
Db 221 GluLysThrTrpSerGlyAspValIleAenLysIleValHisHisThrSerPheAspVal 240
Qy 723 -----GACCAGTGTGCAAGCTGAGCGCTGAGCCCTGCGCC----- 752
Db 241 MetLysAspAenProMetAlaAenHisThrAlaValProAlaHisIlePheAenHisSer 260
Qy 753 -----GTGGCGCGGGAAGAGTTGGGCTGTGGAAGGACATCTTCAACGCTCTCC 800
Db 261 IleSerLysPheMetArgLysGlyMetProGlyAspTrpLysAenHisPheThrValala 280
Qy 801 ATGAATGAGAAGTTTGACTTGGTGTATAAACAGAAAGATGGAAAAGTGTGACCTCACGTTT 860
Db 281 LeuAenGluAenPheAspLysHisTyrGluLysLysMetAlaGlySerThrLeuAenPhe 300

```

RESULT 57

```

US-10-072-012-582
; Sequence 582, Application US/10072012
; Publication No. US2004003493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie

```

APPLICANT: Taupier Jr, Raymond J.
APPLICANT: Gusev, Vladimir Y.
APPLICANT: Colman, Steven D.
APPLICANT: Wolenc, Adam R.
APPLICANT: Pena, Carol E. A
APPLICANT: Furtak, Katarzyna
APPLICANT: Grosse, William M.
APPLICANT: Alsobrook II, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-258
CURRENT APPLICATION NUMBER: US/10/072,012
CURRENT FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: 60/265,102
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/265,514
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,517
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,412
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,395
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/266,406
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/266,767
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: 60/267,057
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/266,975
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/267,459
PRIOR FILING DATE: 2001-02-08
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1391
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 582
LENGTH: 304
TYPE: PRT
ORGANISM: Homo sapiens
US-10-072-012-582

Alignment Scores:
Pred. No.: 1.81e-24 Length: 304
Score: 401.00 Matches: 93
Percent Similarity: 51.7% Conservative: 62
Best Local Similarity: 31.0% Mismatches: 125
Query Match: 9.1% Indels: 20
Gaps: 5

US-10-768-158-1 (1-2419) x US-10-072-012-582 (1-304)

QY 21 ATGGCGGAGCGCGGAGACCCCGGAGGAGTTTCGAG---AGCAAGTAC 77
Db 1 MetAlaLysIleGluLysAsnAlaProThrMetGluLysLysProGluLeuPheAsnIle 20

QY 78 TTCGAGTTCCATGGCGTGGCGGCGCGCCCTTCGCGCGGGAAGATGGAGGAGATCGCC 137
Db 21 MetGluValAspGlyValProThrLeuLeuLeuSerLysGluTrpTrpGluLysValCys 40

QY 138 AACTTCCCGTGGCGGCGGAGCGACGTGGATCGTACCTACCCCAAGTCCGGACCGACG 197
Db 41 AsnPheGlnAlaLysProAspAspLeuAlaThrTyProLysSerGlyThrThr 60

QY 198 TTGCTGCAGGAGTGGTCTACTTGTGTGAGCCGCGGCGTACCCCGATGAGATCGGCTTG 257
Db 61 TrpMetHisGluIleLeuAspMetIleLeuAsnAspGlyAspValGluLysCysLysArg 80

QY 258 ATGAACATCGACGACGCTCCGCTCTGGAG-----TACCCA-----CAGCCG 302
Db 81 AlaGlnThrLeuAspArgHisAlaPheLeuGluLeuLysPheProHisLysGluLysPro 100

QY 303 GGCTTGACATCATCAAGGAACCTGACCTCTCCCGCCTCATCAAGAGCCACCTGCGCTAC 362
Db 101 AspLeuGluPheValLeuGluMetSerProGlnLeuIleLysThrHisLeuProSer 120

QY 363 CGCTTCTGCGCTCTGACCTCCACAATGGAGACTCCCAAGGTCTATATATGGCTCGCAAC 422
Db 121 HisLeuIleProProSerIleTrpLysGluAsnCysLysIleValTrpValAlaArgAsn 140

QY 423 CCAAGGATCTGGTGGTCTTATATACGTTCCACCGCTCTCTGCGGACATGAGGTAC 482
Db 141 ProlysAspCysLeuValSerTyTrHisPheHisArgMetAlaSerPheMetProAsp 160

QY 483 CGAGGCACCTTCAAGGATTTCCCGGAGGCTTATGAATGATAGCTGGGCTACGGCTCC 542
Db 161 ProGlnAsnLeuGluGluPheTyTrpLysPheMetSerGlyLysValValGlyGlySer 180

QY 543 TGGTTTGACGACGTCGAGGAGTTCTGGGAGCACCGCATGGACTCGAAGCTGCTTTTCTC 602
Db 181 TrpPheAspHisValLysGlyTrpTrpAlaAlaLysAspMethHisArgIleLeuTyLeu 200

QY 603 AAGTATGAACATGACATCGGACCTGGTGACGATGGTGAGGAGCTGGCCAGATTCTGTG 662
Db 201 PheTyTrpGluAspIleLysLysAsnProLysHisGluIleHisLysValLeuGluPheLeu 220

QY 663 GGGGTGCTCTGTGACAGGCCCGAGCTGGAAGCCCTGACGAGGAGCTGCCACCGCTGTG 722
Db 221 GluLysThrTrpSerGlyAspValIleAsnLysIleValHisHisThrSerPheAspVal 240

QY 723 -----GACCAGTCTGCAACGCTGAGGCGCTGCCC----- 752
Db 241 MetLysAspAsnProMetAlaAsnHisThrAlaValProAlaHisIlePheAsnHisSer 260

QY 753 -----GTGGCGCGGGAAGAGTTGGGCTGTGGAAGGACATCTTACCGCTCTCC 800
Db 261 IleSerLysPheMetArgLysGlyMetProGlyAspTrpLysAsnHisPheThrValAla 280

QY 801 ATGATGAGAAGTTTGACTTGGTGTATATAACAGAGATGGGAAGTGTGACCTCACCTTT 860
Db 281 LeuAsnGluAsnPheAspLysHisTyTrpGluLysMetAlaGlySerThrLeuAsnPhe 300

RESULT 58
US-11-108-875-4
; Sequence 4, Application US/11108875
; Publication No. US20050196802A1
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-4
; CURRENT APPLICATION NUMBER: US/11/108,875
; CURRENT FILING DATE: 2005-04-19
; PRIOR APPLICATION NUMBER: 10/199,334
; PRIOR FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-11-108-875-4

Alignment Scores:
Pred. No.: 1.81e-24 Length: 304
Score: 401.00 Matches: 93
Percent Similarity: 51.7% Conservative: 62

QY 483 CGAGGCACCTTCAAGATCTCGCGGAGGTTTATGATGATAGCTGGGCTACGGCTCC 542
 Db 161 ProGlnAsnLeuGluPheTyrGluLysPheMetSerGlyLysValValGlyArgSer 180
 QY 543 TGGTTTGAGCAGCTGCGAGGAGTCTCGGAGCACCGCATGGACTCGAACGTCGCTTTTCTC 602
 Db 181 TrpPheAspHisValLysGlyTrpTrpAlaAlaLysAspThrHisArgIleLeuTyrLeu 200
 QY 603 AAGTATGAAGACATGTCATCGGACCTGCTGACGATGGTGAGCAGCTGGCCAGATTCTCTG 662
 Db 201 PheTyrGluAspIleLysLysAsnProLysHisGluIleHisLysValLeuGluPheLeu 220
 QY 663 GGGGTGCTCGTGACAAAGCCACCTGGAAGCCCTGAGGAGCACTGCCACCAGCTGGTG 722
 Db 221 GluLysThrLeuSerGlyAspValIleAsnLysIleValHisHisThrSerPheAspVal 240
 QY 723 -----GACCAGTGTGCAAGCTGAGGCCCTGGCC----- 752
 Db 241 MetLysAspAsnProMetAlaAsnHisThrAlaValProAlaHisIlePheAsnHisSer 260
 QY 753 -----GTGGCCGGGGAAGAGTTGGGCTGTGGAAGGACATCTTCACCGTCTCC 800
 Db 261 IleSerLysPheMetArgLysGlyMetProGlyAspTrpLysAsnHisPheThrValAla 280
 QY 801 ATGAATGAGAAGTTTGACTGCTGTATAAACAGAAAGATGGAAAGCTGACCTCACGTTT 860
 Db 281 MetAsnGluAsnPheAspLysHisTyrGluLysLysMetAlaGlySerThrLeuAsnPhe 300

RESULT 60

US-11-079-743-9
 ; Sequence 9, Application US/11079743
 ; Publication No. US20050181415A1
 ; GENERAL INFORMATION:
 ; APPLICANT: INCYTE GENOMICS, INC.
 ; APPLICANT: YANG, Junming
 ; APPLICANT: BAUGHN, Mariah R.
 ; APPLICANT: BURFORD, Neil
 ; APPLICANT: AU-YOUNG, Janice
 ; APPLICANT: LU, Dyoung Aina M.
 ; APPLICANT: REDDY, Roopa
 ; APPLICANT: RING, Huijun Z.
 ; APPLICANT: HILLMAN, Jennifer L.
 ; APPLICANT: YUE, Henry
 ; APPLICANT: AZIMZAI, Yalda
 ; APPLICANT: YAO, Monique G.
 ; APPLICANT: GANDHI, Ameena R.
 ; APPLICANT: NGUYEN, Denniel B.
 ; APPLICANT: BAUGHN, Mariah R.
 ; APPLICANT: TANG, Y. Tom
 ; APPLICANT: LAL, Preeti
 ; APPLICANT: YUE, Henry
 ; APPLICANT: BANDMAN, Olga
 ; TITLE OF INVENTION: DRUG METABOLIZING ENZYMES
 ; FILE REFERENCE: PI-0007 PCT
 ; CURRENT APPLICATION NUMBER: US/11/079,743
 ; CURRENT FILING DATE: 2005-03-14
 ; PRIOR APPLICATION NUMBER: US/10/181,108
 ; PRIOR FILING DATE: 2002-07-11
 ; PRIOR APPLICATION NUMBER: US/10/181,108
 ; PRIOR FILING DATE: 2002-07-11
 ; NUMBER OF SEQ ID NOS: 48
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 9
 ; LENGTH: 304
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: Incyte ID No: 638419CD1
 ; US-11-079-743-9

Alignment Scores:

Pred. No.: 2,67e-24 Length: 304
 Score: 399.00 Matches: 93
 Percent Similarity: 51.3% Conservatives: 61
 Best Local Similarity: 31.0% Mismatches: 126
 Query Match: 9.1% Indels: 20
 DB: 6 Gaps: 5

US-10-768-158-1 (1-2419) x US-11-079-743-9 (1-304)

QY 21 ATGCGGAGAGCGAGCGAGACCCGACCCAGCCCGGGGAGTTCGAG---AGCAAGTAC 77
 Db 1 MetAlaLysIleGluLysAsnAlaProThrMetGluLysLysProGluLeuPheAsnIle 20
 QY 78 TTCGAGTTCATCGCGTGGCTGCCGCCCTTCGCGCGGGAAGATGGAGGAGATGCC 137
 Db 21 MetGluValAspGlyValProThrLeuIleLeuSerLysGluTrpTrpGluLysValCys 40
 QY 138 AACTTCCCGGTGGCGCCAGCGACGTGTGATCGTCACCTACCCCAAGTCCGGCACCAGC 197
 Db 41 AsnPheGlnAlaLysProAspAspLeuIleLeuAlaThrTyrProLysSerGlyThrThr 60
 QY 198 TTCTGTCAGGAGTGTCTACTTTGGTGAGCCAGCGGCTGACCCCGATGAGATCGGCTG 257
 Db 61 TrpMetHisGluIleLeuAspMetIleLeuAsnAspGlyAspValGluLysCysLysArg 80
 QY 258 ATGAACATCGACGAGCTCCCGGTCTCTGGAG-----TACCCA-----CAGCCG 302
 Db 81 AlaGlnThrLeuAspArgHisAlaPheLeuGluLysPheProHisLysGluLysPro 100
 QY 303 GGCTGGACATCATCAAGAACTGACCTCTCCCGCTCATCAAGAGCCACCTGCCCTAC 362
 Db 101 AspLeuGluPheValLeuGluMetSerSerProGlnLeuIleLysThrHisLeuProSer 120
 QY 363 CGTTTCTGCTCTGACCTCCACAATGGAGATCCCAAGTCACTATATGTTGGTTCGCAAC 422
 Db 121 HisLeuIleProProSerIleTrpLysGluAsnCysLysIleValTyrValAlaArgAsn 140
 QY 423 CCAAGGATCTGGTGTCTTATTACGTTCCACCGCTCTCTCGGACCATGAGCTAC 482
 Db 141 ProLysAspCysLeuValSerTyrHisPheHisArgMetAlaSerPheMetProAsp 160
 QY 483 CGAGGCACCTTTCAGAAATTCGCGGAGGTTTATGAATGATTAAGCTGGGCTACGGCTCC 542
 Db 161 ProGlnAsnLeuGluPheTyrGluLysPheMetSerGlyLysValValGlyArgSer 180
 QY 543 TGGTTGAGCAGCTGCGAGGAGTCTGGGAGCACCGCATGGACTCGCAACGCTTTTCTC 602
 Db 181 TrpPheAspHisValLysGlyTrpTrpAlaAlaLysAspThrHisArgIleLeuTyrLeu 200
 QY 603 AAGTATGAAGACATGTCATCGGACCTGGTGACGATGGTGAGCAGCTGGCCAGATTCTCTG 662
 Db 201 PheTyrGluAspIleLysLysAsnProLysHisGluIleHisLysValLeuGluPheLeu 220
 QY 663 GGGGTGCTCGTGACAAAGCCAGCTGGAAGCCCTGACGAGCACTGCCACCAGCTGGTG 722
 Db 221 GluLysThrLeuSerGlyAspValIleAsnLysIleValHisHisThrSerPheAspVal 240
 QY 723 -----GACCAGTGTCTCAACGCTGAGGCCCTGGCC----- 752
 Db 241 MetLysAspAsnProMetAlaAsnHisThrAlaValProAlaHisIlePheAsnHisSer 260
 QY 753 -----GTGGCCGGGGAAGAGTTGGGCTGTGGAAGGACATCTTCACCGTCTCC 800
 Db 261 IleSerLysPheMetArgLysGlyMetProGlyAspTrpLysAsnHisPheThrValAla 280
 QY 801 ATGAATGAGAAGTTTGACTGCTGTATAAACAGAAAGATGGAAAGCTGACCTCACGTTT 860
 Db 281 MetAsnGluAsnPheAspLysHisTyrGluLysLysMetAlaGlySerThrLeuAsnPhe 300

RESULT 61

US-10-072-012-206
 ; Sequence 206, Application US/10072012
 ; Publication No. US20040033493A1

```

; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meeta
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072.012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 206
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-072-012-206

Alignment Scores:
Pred. No.: 3,53e-24 Length: 295
Score: 397.50 Matches: 85
Percent Similarity: 54.2% Conservative: 58
Best Local Similarity: 32.2% Mismatches: 102
Query Match: 9.0% Indels: 19
DB: 4 Gaps: 4

US-10-768-158-1 (1-2419) x US-10-072-012-206 (1-295)
Qy 126 GAGGATCGCAACTCCGGTGGCGCCAGCAGCGTGTGGATCGTCACCTACCCCAAG 185
Dy |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Dy 28 GluysValCysAsnPhelAlaLysProAspLeuIleLeuAlaThyTrpLys 47
Qy 186 TCCGACACAGCTGCTCAGAGGTGTCTACTTGTGGACCGAGCGGCGCTGACCCCGAT 245
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

Db 48 SerGlyThrTrpMetHisGluIleLeuAspMetIleLeuAsnAspGlyAspValGlu 67

Qy 246 GAGATCGGCTTGATGAACATCGACAGCAGCTCCCGTCTCTGGAG-----TACCCA--- 296

Db 68 LysCysLysArgAlaGlnThrLeuAspArgHisAlaPheLeuGluLeuLysPheProHis 87

Qy 297 -----CAGCGGGCTGGACATCAAGAACTGACCTCTCCCGCTCATCAAGAGC 350

Db 88 LysGluLysProAspLeuGluPheValLeuGluMetSerSerProGlnLeuIleLysThr 107

Qy 351 CACCTGCCCTACCGCTTCTGCGCTCTGACCTCACAATGGAGACTCCAAGGTCTAT 410

Db 108 HisLeuProSerHisLeuIleProProSerIleTrpLysGluAsnCysLysIleValTr 127

Qy 411 ATGGCTCGCAACCCCAAGATCTGGTGTCTTATTTATCAGTTCCACCGCTCTCTCGG 470

Db 128 ValAlaArgAsnProLysAspCysLeuValSerTrpTrpHisPheHisArgMetAlaSer 147

Qy 471 ACCATGAGCTACCGAGGCACCTTTCAGAATTCGCGGAGGTTTATGAATGATAAGCTG 530

Db 148 PheMetProAspProGlnAsnLeuGluGluPheTrpGluLysPheMetSerGlyLysVal 167

Qy 531 GGCTACGGCTCTGGTTTGGACGCTGACGAGTCTCGGGAGCACCCGATGGACTCGAAC 590

Db 168 ValGlyGlySerTrpPheAspHisMetLysGlyTrpTrpAlaAlaLysAspMetHisArg 187

Qy 591 GTGCTTTTCTCAAGTATGAAGACATGCATCGGACCTGTGTGACGATGGTGGAGCAGCTG 650

Db 188 IleLeuTrpLeuPheTrpGluAspIleLysLysAsnProLysHisGluIleHisLysVal 207

Qy 651 GCCAGATTCTGGGGTGTCTGTGACAAAGCCCGCTGGAAGCCCTGACGGAGCAGCTGC 710

Db 208 LeuGluPheLeuGluLysThrTrpSerGlyAspValIleAsnLysIleValHisHisThr 227

Qy 711 CACCAGCTGTGT-----GACCAGCTGTGCAACGCTGAGGCCCTGCC----- 752

Db 228 SerPheAspValMetLysAspAsnProMetAlaAsnHisThrAlaValProAlaHisIle 247

Qy 753 -----GTGGCGCGGGAAGAGTTGGGCTGTGGAAGACATC 788

Db 248 PheAsnHisSerIleSerLysPheMetArgLysGlyMetProGlyAspTrpLysAsnHis 267

Qy 789 TTACCGCTTCCATGAATGAGAAGTTTTCAGCTGTGTATATAACAGAGATGGGAAGTGT 848

Db 268 PheThrValAlaMetAsnGluAsnPheAspLysHisTrpGluLysLysMetAlaGlySer 287

Qy 849 GACCTCACGTTT 860

Db 288 ThrLeuAsnPhe 291

RESULT 62

US-09-898-570-26

; Sequence 26, Application US/09898570

; Patent No. US20020123612A1

; GENERAL INFORMATION:

; APPLICANT: GERLACH, VALERIE L.

; APPLICANT: ELLERMAN, KAREN

; APPLICANT: MACDOUGALL, JOHN R.

; APPLICANT: SMITHSON, GLENDA

; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND

; FILE REFERENCE: 15966-776CIP

; CURRENT APPLICATION NUMBER: US/09/898,570

; CURRENT FILING DATE: 2001-07-03

; PRIOR APPLICATION NUMBER: 60/198,293

; PRIOR FILING DATE: 2000-04-19

; PRIOR APPLICATION NUMBER: 60/198,645

; PRIOR FILING DATE: 2000-04-20

; PRIOR APPLICATION NUMBER: 60/210,809

; PRIOR FILING DATE: 2000-06-09

; PRIOR APPLICATION NUMBER: 60/199,476

; PRIOR FILING DATE: 2000-04-26

; PRIOR APPLICATION NUMBER: 60/200,025

```
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/224,610
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: 60/200,024
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/199,880
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/218,591
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 60/271,814
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/215,855
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 09/839,446
; PRIOR FILING DATE: 2001-04-19
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 305
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: POLYX
; OTHER INFORMATION: h_nh0443k08_A
US-09-898-570-26

Alignment Scores:
Pred. No.: 4,34e-24 Length: 305
Score: 396.50 Matches: 93
Percent Similarity: 50.8% Conservative: 60
Best Local Similarity: 30.9% Mismatches: 127
Query Match: 9.0% Indels: 21
DB: 3 Gaps: 5

US-10-768-158-1 (1-2419) x US-09-898-570-26 (1-305)
QY 21 ATGCGGAGAGCGAGCGGAGACCCGAGACCCCGGGGGAGTTCGAG---AGCAAGTAC 77
Db 1 MetAlaLysLeuGluLysAsnAlaProThrMetGluLysLysProGluLeuPheAsnIle 20
QY 78 TTCGAGTTCCATGCGTGGCGGCTCCGCGCCCTTCTGCGCGGGAAGAGATGAGATCGCC 137
Db 21 MetGluValAlaAspGlyValProThrLeuLeuSerLysGluLysGluLysValCys 40
QY 138 AACTTCCGGTGGCGGCGGAGCGAGCGTGTGGATCTCACCACCCAGTCCGCGGACGAGC 197
Db 41 AsnPheGlnAlaLysProAspAspLeuLeuAlaThrTyrProLysSerGlyThrThr 60
QY 198 TTGCTGAGGAGGTTGCTACTTGTGAGCGGCGTGTGACCCCGGATGAGATCGGCTTG 257
Db 61 TrpMetHisGluLeuLeuAspMetIleLeuAsnAspGlyAspValGluLysCysLysArg 80
QY 258 ATGAACATCGACGAGCGTCCCGTCTCGAG-----TACCCA-----CAGCCG 302
Db 81 AlaGlnThrLeuAspArgHisAlaPheLeuGluLysPheProHisLysGluLysPro 100
QY 303 GGCTTGACATCATCAGGAAGTCACTCTCCCGCTCATCAGAGCCAGCTCCCTCCCTAC 362
Db 101 AspLeuGluPheValLeuGluMetSerSerProGlnLeuLeuLysThrHisLysProSer 120
QY 363 CGCTTTCTGCCTCTGACCTCCACAATGGAGACTCCCAAGGTCACTATATGGCTCGCAAC 422
Db 121 HisLeuLeuProProSerIleTrpLysGluAsnCysLysIleValTyrValAlaArgAsn 140
QY 423 CCCAAGGATCTGGTGTCTTATATATCAGTTCCACCGCTCTCTCGGACCATGAGCTAC 482
Db 141 ProlLysAspCysLeuValSerTyrTyrHisPheHisArgMetAlaSerPheMetProAsp 160
QY 483 CGAGGACCTTTCAGAAATCTCGCGGAGGTTTATGATGATAGCTGGGCTACGGCTCC 542
Db 161 ProGlnAsnLeuGluPheTyrGluLysPheMetSerGlyLysGlyGluPheGlySer 180
QY 543 TGGTTTGAGCAGCGTGCAGGAGTTCTGGGAGCAGCGCATGGAAGTGGCTTTTCTC 602
```

```
Db 181 TrpPheAspHisValLysGlyTrpTrpAlaAlaLysAspMethHisArgIleLeuTyrLeu 200
QY 603 AAGTATGAGACATGTCATCGGACCTGCTGACGATGCTGAGCAGCTGGCCAGATTCTGTG 662
Db 201 PheTyrGluAspIleLysGlnAsnProLysHisGluIleHisLysValLeuGluPheLeu 220
QY 663 GGGGTGCTCTGTGACAAAGGCCAGCTGGAAGCCCTGACGAGCAGCTGCCACGAGCTGGT 722
Db 221 GluLysThrTrpSerGlyAspValIleAsnLysIleValHisThrSerPheAspVal 240
QY 723 -----GACCAGTCTGCAACGCTGAGCGCTGCCCTG----- 755
Db 241 MetLysAspAsnProMetAlaAsnHisThrAlaValProAlaHisIlePheAsnHisSer 260
QY 756 -----GCCCGGGGAAGATTGGCTGTGGAAGGACATCTTCCACGTC 797
Db 261 IleSerLysPheMetArgLysGlyGlyMetProGlyAspTrpLysAsnHisPheThrVal 280
QY 798 TCCATGAATGAGAAGTTTGACTTGGTGTATATAACAGAGATGGAAAGTGTGACCTCACG 857
Db 281 AlaLeuAsnGluAsnPheAspLysHisTyrGluLysLysMetAlaGlySerThrLeuAsn 300
QY 858 TTT 860
Db 301 Phe 301

RESULT 63
US-09-839-446-26
; Sequence 26, Application US/09839446
; Publication No. US20030050232A1
; GENERAL INFORMATION:
; APPLICANT: GERLACH, VALERIE L.
; APPLICANT: ELLERMAN, KAREN
; APPLICANT: MACDOUGALL, JOHN R.
; APPLICANT: SMITHSON, GLENDA
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND
; TITLE OF INVENTION: METHODS OF USING THE SAME
; FILE REFERENCE: 15966-776
; CURRENT APPLICATION NUMBER: US/09/839,446
; CURRENT FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: 60/198,293
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: 60/198,645
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: 60/210,809
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/199,476
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/200,025
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/224,610
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: 60/200,024
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/199,880
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/218,591
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 60/271,814
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 305
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: POLYX
; OTHER INFORMATION: h_nh0443k08_A
US-09-839-446-26

Alignment Scores:
```


;
; PRIOR APPLICATION NUMBER: 60/249,216
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,210
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/226,681
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,759
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/225,213
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/227,182
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,214
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/235,836
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/230,438
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/215,135
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/225,266
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/249,218
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,208
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,213
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,212
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,207
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,245
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,244
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,217
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,211
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,215
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,264
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,214
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,297
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/232,400
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/231,242
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,081
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,080
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,414
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,244
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/233,064
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/233,083
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,397
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,399
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,401
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/241,808
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,826

;
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,786
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,221
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,475
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/231,243
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/233,065
; PRIOR FILING DATE: 2000-09-14

Alignment Scores:
Pred. No.: 4.28e-23 Length: 76
Score: 382.00 Matches: 71
Percent Similarity: 97.3% Conservative: 1
Best Local Similarity: 95.9% Mismatches: 2
Query Match: 8.7% Indels: 0
DB: 3 Gaps: 0

US-10-768-158-1 (1-2419) x US-09-989-442-149 (1-76)
Qy 288 GAGTACCCACAGCGGGCTGGACATCATCAAGGAACCTCTCCCGCCTCATCAAG 347
Db 1 LysTyrProGlnProGlyLeuAspIleIle**GlulLeuThrSerProArgLeuIleLys 20
Qy 348 AGCCACCTGCCCTACGGCTTTCGCTCTGACCTCCACAAATGGAGACTCCAAGGTGATC 407
Db 21 SerHisLeuProTyrArgPheLeuProSerAspLeuHisAsnGlyAspSerLysValIle 40
Qy 408 TATATGGCTCGCAACCCCAAGGATCTGGTGTCTTATTATCAGTTCCACCGCTCTCTG 467
Db 41 TyrMetAlaArgAsnProLysAspLeuValSerTyrTyrGlnPheHisArgSerLeu 60
Qy 468 CGGACCATGAGTACCGAGGACCTTTCAAGAATTCGCGG 509
Db 61 ArgThrMetSerTyrArgGly***PheGlnGluPheCysArg 74

RESULT 67
US-10-756-149-5061
; Sequence 5061, Application US/10756149
; Publication No. US20050181375A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
; FILE REFERENCE: file
; CURRENT APPLICATION NUMBER: US/10/756,149
; CURRENT FILING DATE: 2004-01-12
; NUMBER OF SEQ ID NOS: 5818
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5061
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-756-149-5061

Alignment Scores:
Pred. No.: 9.22e-23 Length: 350
Score: 381.00 Matches: 94
Percent Similarity: 51.8% Conservative: 52
Best Local Similarity: 33.3% Mismatches: 108
Query Match: 8.7% Indels: 28
DB: 5 Gaps: 7

US-10-768-158-1 (1-2419) x US-10-756-149-5061 (1-350)
Qy 54 CCGGGGAGGTTTCGAGACAGTACTTCGAGTTCCATCGGCTGGCGGCTGCCCGCTTCTGC 113
Db 13 ProGlyGlu-----TyrPheArgTyrLysGlyVal-----ProPhePro 25
Qy 114 CGCGGG-----AAGATGGAGGAGATCGCC-----AACTTCCCGTGGCGGCC 155

```
Db 26 ValGlyLeuYrSerLeuGluSerLeuLeuAlaGluAenThrGlnAspValArgAsp 45
Qy 156 AGCGAGGTGCGATCGTACCTACCCCAAGTCCGCGACACGCTTCTCGTCCAGGAGTGTCTC 215
Db 46 AspAspIlePheIleIleThrTrpProLysSerGlyThrTrpMetIleGluIleIle 65
Qy 216 TACTGTGTGACCGACGGCGCTGACCCCGATGAGATCGCTTTGATGAACATCGACGAGCAG 275
Db 66 CysLeuIleLeuLysGluGlyAspProSerTrpIleArgSerValProIleTrpGluArg 85
Qy 276 CTCCGGTCTCGGAGTACCCACGCGGCTGGACATCATCAAGGAACCTACCTCCCTCCC 335
Db 86 AlaProTrpCysGlu--ThrIleValGlyAlaPheSerLeuProAspGlnTrpSerPro 104
Qy 336 CGCTCTCATCAAGACCGCTACCGCTTCTGCGCTTCTGCGCTCCACATCCCAATGGAGAC 395
Db 105 ArgLeuMetSerSerHisLeuProIleGlnIlePheThrLysAlaPhePheSerSerLys 124
Qy 396 TCCAAGTCTATATATGCTCGCAACCCCAAGGATCGTGGTGTCTTATATCATGATTC 455
Db 125 AlalysValIleIleYrMetGlyArgAsnProArgAspValValSerLeuYrHisTrp 144
Qy 456 CACGCTCTCGCGGACCATGAGTACCGAGGACCTTTCAAGAAATTCGCGGAGGTTT 515
Db 145 SerLysIleAlaGlyGlnLeuLysAspProGlyThrProAspGlnPheLeuArgAspPhe 164
Qy 516 ATGAATGATAGTGGGTACGGCTCCGCTGCTTGTGACGACGTCGAGGAGTCTGGGAGCAC 575
Db 165 LeuLysGlyGluValGlnPheGlySerTrpPheAspHisIleLysGlyTrpLeuArgMet 184
Qy 576 CGCATGGAAGTCTCGAAGTCTTTTCTCAAGTATGAAGACATGCATCGGAGCTTGTGACG 635
Db 185 LysGlyLysAspAsnPheLeuPheIleThrTrpGluGluLeuGlnAspLeuGlnGly 204
Qy 636 ATGTTGAGCAGCTGGCCAGATTCCTGGGGTGTCTGTGACAGGCGCCACCTGGAAGCC 695
Db 205 SerValGluArgIleCysGlyPheLeuGlyArgProLeuGlyLysGluAlaLeuGlySer 224
Qy 696 CTGACGGAGCAC-----TGCACACAGCTGTGTGGACCAAGTGTGCAACGCTGAGGCC 746
Db 225 ValValAlaHisSerThrPheSerAlaMetLysAlaAenThrMetSerAsnTrpThrLeu 244
Qy 747 CTGCCC-----GTGGCGCGGGGAAGAGTTGGG 773
Db 245 LeuProSerLeuLeuAspHisArgGlyAlaPheLeuArgLysGlyValCysGly 264
Qy 774 CTGTGGAGGACATCTTCACCGTCTCCATGAATGAGAAGTTGATGGTGTATATAACAG 833
Db 265 AspTrpLysAsnHisPheThrValAlaGlnSerGluAlaPheAspArgAlaTrpArgLys 284
Qy 834 AAGATG 839
Db 285 GlnMet 286

RESULT 68
US-11-097-143-11343
; Sequence 11343, Application US/11097143
; Publication No. US2005020858A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CLO00728
; CURRENT FILING DATE: 2005-04-04
; PRIOR FILING DATE: 1999-10-05
; PRIOR FILING DATE: 1999-10-19
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
```

```
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11343
; LENGTH: 338
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-11343

Alignment Scores:
Pred. No.: 2,7e-21 Length: 338
Score: 363.50 Matches: 91
Percent Similarity: 48.3% Conservative: 53
Best Local Similarity: 30.5% Mismatches: 119
Query Match: 8.3% Indels: 35
DB: 6 Gaps: 7

US-10-768-158-1 (1-2419) x US-11-097-143-11343 (1-338)
Qy 12 GGCGCGCGCATGGCGGAGAGGAGGCGGAGACCCCGACCCCGCG-----GGGAGTTC 65
Db 23 GlyGlyYrSerSerIlePheAlaSerSerLysProSerValProValValGlyAsnTrp 42
Qy 66 GAGAGCAAGTACTTCGAGTTCATGCGGTGCGCGCTCGCGCTTCTGCGCGGGAAGATG 125
Db 43 GluGlnArgPhe-----CysArgLeuAlaAspThrPheGlnProValLeu 57
Qy 126 GAGGAGATCCGCCAATTCCCGGTGCGCGCCAGCAGCGTGTGGATCGTCACTACCCCAAG 185
Db 58 AspArgValTrpAspPheGluValArgAspAspValTrpIleValThrLeuProLys 77
Qy 186 TCGCGGACACAGCTGTGTCGAGGAGTGTCTACTTGTGTGAGCCAGGGCGCTGACCCCGAT 245
Db 78 CysGlyThrTrpMetGlnGluLeuAlaTrpLeuValIleAsnGluCysAspPheGlu 97
Qy 246 GAGATCGGCTTGATGAACATCGACGACGAGCTCCCGTCTCTGGAGTAC----- 293
Db 98 ThrAlaLysSerValAspLeuThrHisArgSerProPheLeuGluPheAsnGlyValVal 117
Qy 294 -----CCACAGCGCGGCGCTGACATCATCAAGGAACCTGACCTCTCCCGCTCATC 344
Db 118 ProAsnValProHisAspThrIleAlaAlaAsnAlaLeuProSerProArgLeuIle 137
Qy 345 AAGAGCCACCTGCGCTACCGCTTCTGCGCTCTGACCTCCACATGGAGACTCCAAGTC 404
Db 138 LysSerHisLeuProAlaTrpMetLeuProArgGlnIleTrpSerLysArgProLysIle 157
Qy 405 ATCTATATGCTCGCAACCCCAAGGATCTGGTGTGTCTTATTATTCAGTTTCCACGCTCT 464
Db 158 IleTrpValTrpArgAsnProLysAspAlaIleSerTrpPheHisHisTrpArgGly 177
Qy 465 CTGCGGACCATGAGTACCGAGGACCTTTCAAGATTTCTGCCGAGAGTTTATGAATGAT 524
Db 178 Met-----ValGlyTrpGlnGlyThrLysSerAspPheMetHisSerPheIleAspGly 195
Qy 525 AAGCTGGCTACGGCTCTCGTTTGACACGTGCGAGGAGTTCTGGGAGCCCGCATGAC 584
Db 196 TyrValAsnPheThrProCysTrpProHisIleLeuAspPheTrpGlnLeuArgHisGlu 215
Qy 585 TCGAAGCTGCTTTTCTCAAGATGAAGACATGATCGGAGACCTGCGTGTGAGTGTGGAG 644
Db 216 ProAsnIlePhePheThrSerTrpGluArgMetLysGlyGlnLeuGlyGlnValIleSer 235
```



```
Qy 84 TTCCATGCGCTGCGGCTCCGCC-----TTCTGCCCGGGGAAGATGGAGGAGATCGCC 137
Db 8 PheGluGlyIleAlaPheProThrMetGlyPheArgSerGluThrLeuArgLysValArg 27
Qy 138 AAC---TTCCCGGTCGCGCCAGCAGCTGCGTGGATCGTCACCTACCCCAAGTCGGGCACC 194
Db 28 AspGluPheValIleArgAspGluAspValIleIleLeuThrTyrProLysSerGlyThr 47
Qy 195 AGCTTGCTGCAGGAGGTGCTACTTGTGTGAGCCAGGCGGTGACCCCGATGAGATCGGC 254
Db 48 AsnTrpLeuAlaGluLeuCysLeuMethHisSerLysGlyAspAlaLysTrpIleGln 67
Qy 255 TTGATGAACATGACGACAGCTCCGCTCTGGAGTACCCACAGCCGGGCTGGACATC 314
Db 68 SerValProIleTrpGluArgSerProTrpValGlu---SerGluIleGlyThrAla 86
Qy 315 ATCAAGGAACCTGACCTCTCCCGCTCATCAAGAGCCACCTACCGCTTCTGCGCC 374
Db 87 LeuSerGluThrGluSerProArgLysPheSerHisLeuProIleGlnLeuPhePro 106
Qy 375 TCTGACCTCCAAATGGAGACTCCAAAGTCTATATGGCTCGCAACCCCAAGGATCTG 434
Db 107 LysSerPhePheSerSerLysAlaLysValIleTyrLeuMetArgAsnProArgAspVal 126
Qy 435 GTGGTGCTTATTATACGTTCCACCGCTCTCTGCGGACCATGAGCTACCGAGCACCTTT 494
Db 127 LeuValSerGlyTyrPhePheTrpLysAsnMetLysPheIleLysProLysSerTrp 146
Qy 495 CAAGAATCTCCCGGACCTTATGATGATGAAGCTGAGCTGCGGCTGCGGCTGCGGCTG 554
Db 147 GluGluTyrPheGluTrpPheCysGlnGlyThrValLeuTyrGlySerTrpPheAspHis 166
Qy 675 GACAAGGCCAGCTGGAAGCCCTGACGAGCACATGCG-----CACCAGCTGGTGAC 725
Db 207 GluProGluGluLeuAsnLeuIleLysAsnSerSerPheGlnSerMetLysGluAsn 226
Qy 759 ---CGGGGAAGAGTTGGGCTGTGGAAGGACATCTTACCGCTCCATGGAATGAGAAGTTT 815
Db 247 ArgLysGlyValSerGlyAspTrpLysAsnHisPheThrValAlaGlnAlaGluAspPhe 266
Qy 816 GACTTGGTGATAACAGAGAATG 839
Db 267 AspLysLeuPheGlnGluLysMet 274
RESULT 71
US-10-851-921-12
; Sequence 12, Application US/10851921
; Publication No. US20040241737A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Khawaja, Xavier
; APPLICANT: Xu, Jun
; APPLICANT: Liang, JinJun
; FILE OF INVENTION: METHODS FOR DIAGNOSING MOOD DISORDERS
; FILE REFERENCE: AM101227
; CURRENT APPLICATION NUMBER: US/10/851,921
; CURRENT FILING DATE: 2004-05-21
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
```

```
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-851-921-12
Alignment Scores: 1.19e-20 Length: 285
Score: 355.50 Matches: 83
Percent Similarity: 53.0% Conservative: 59
Best Local Similarity: 31.0% Mismatches: 109
Query Match: 8.1% Indels: 17
DB: 5 Gaps: 5
US-10-768-158-1 (1-2419) x US-10-851-921-12 (1-285)
Qy 84 TTCCATGCGCTGCGGCTCCGCC-----TTCTGCCCGGGGAAGATGGAGGAGATCGCC 137
Db 8 PheGluGlyIleAlaPheProThrMetGlyPheArgSerGluThrLeuArgLysValArg 27
Qy 138 AAC---TTCCCGGTCGCGCCAGCAGCTGCGTGGATCGTCACCTACCCCAAGTCGGGCACC 194
Db 28 AspGluPheValIleArgAspGluAspValIleIleLeuThrTyrProLysSerGlyThr 47
Qy 195 AGCTTGCTGCAGGAGGTGCTACTTGTGTGAGCCAGGCGGTGACCCCGATGAGATCGGC 254
Db 48 AsnTrpLeuAlaGluLeuCysLeuMethHisSerLysGlyAspAlaLysTrpIleGln 67
Qy 255 TTGATGAACATGACGACAGCTCCGCTCTGGAGTACCCACAGCCGGGCTGGACATC 314
Db 68 SerValProIleTrpGluArgSerProTrpValGlu---SerGluIleGlyThrAla 86
Qy 315 ATCAAGGAACCTGACCTCTCCCGCTCATCAAGAGCCACCTACCGCTTCTGCGCC 374
Db 87 LeuSerGluThrGluSerProArgLysPheSerHisLeuProIleGlnLeuPhePro 106
Qy 375 TCTGACCTCCAAATGGAGACTCCAAAGTCTATATGGCTCGCAACCCCAAGGATCTG 434
Db 107 LysSerPhePheSerSerLysAlaLysValIleTyrLeuMetArgAsnProArgAspVal 126
Qy 435 GTGGTGCTTATTATACGTTCCACCGCTCTCTGCGGACCATGAGCTACCGAGCACCTTT 494
Db 127 LeuValSerGlyTyrPhePheTrpLysAsnMetLysPheIleLysProLysSerTrp 146
Qy 495 CAAGAATCTCCCGGACCTTATGATGATGAAGCTGAGCTGCGGCTGCGGCTGCGGCTG 554
Db 147 GluGluTyrPheGluTrpPheCysGlnGlyThrValLeuTyrGlySerTrpPheAspHis 166
Qy 555 GTGACGAGGTTCTGGGAGCACCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 614
Db 167 IleHisGlyTrpMetProMetArgGluGluLysAsnPheLeuLeuSerTrpGluGlu 186
Qy 615 ATGATCGGAGGTTCTGGGAGCACCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 674
Db 167 IleHisGlyTrpMetProMetArgGluGluLysAsnPheLeuLeuSerTrpGluGlu 186
Qy 615 ATGATCGGAGGTTCTGGGAGCACCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 674
Db 187 LeuLysGlnAspThrGlyArgThrIleGluLysIleCysGlnPheLeuGlyLysThrLeu 206
Qy 675 GACAAGGCCAGCTGGAAGCCCTGACGAGCACATGCG-----CACCAGCTGGTGAC 725
Db 207 GluProGluGluLeuAsnLeuIleLysAsnSerSerPheGlnSerMetLysGluAsn 226
Qy 726 CAGTGTGCAACGCTGAGGCGCTGCGGCTGGGC----- 758
Db 227 LysMetSerAsnTyrSerLeuLeuSerValValAspTyrValValAspLysAlaGlnLeuLeu 246
Qy 759 ---CGGGGAAGAGTTGGGCTGTGGAAGGACATCTTACCGCTCCATGGAATGAGAAGTTT 815
Db 247 ArgLysGlyValSerGlyAspTrpLysAsnHisPheThrValAlaGlnAlaGluAspPhe 266
Qy 816 GACTTGGTGATAACAGAGAATG 839
Db 267 AspLysLeuPheGlnGluLysMet 274
RESULT 72
US-11-179-843-3
```


Db 109 PhePheLysSerLysAlaLysMetIleTyrIleileArgAsnProArgAspIleileVal 128
Qy 441 TCTTATTATCAGTTCACCGCTCTCTCGGACCATGAGTACCGAGGACACCTTTCAAGAA 500
Db 129 SerGlyTyrPhePheTrpLysSerThrAsnLeuValLysArgProGluSerLeuGluGln 148
Qy 501 TCTGCCGGAGGTTTATGATGATAAGCTGGGCTACGGCTCCTGTTGTGACACGTCGAG 560
Db 149 TyrPheGluTrpPheIleGlnGlyAsnValProTyrGlySerTrpPheAspHisIleArg 168
Qy 561 GAGTCTTGGGAGCACCGCATGAGTCTCGAACGTGCTTTTCTCAAGTATGAGACATGCAT 620
Db 169 GlyTrpLeuProMetArgAspLysGluAsnValLeuLeuLeuSerTyrGluGluLeuLys 198
Qy 621 CGGACACCTGGTGACGATGGTGCGGACGAGTCCCGCATTCCTCGGGGGTCTCCTGTGCAAG 680
Db 189 ArgAspThrArgSerAlaValGluLysIleCysGlnPheLeuGlyLysLeuGluPro 208
Qy 681 GCCCAGCTGGAGCCCTGACGAGCAGTGCACGAGCTGGTGAGCCAGTGTGTCGAAC--- 737
Db 209 GluGluLeuSerSerValValGluAsnSerSerPheGlnValMetLysGluAsnMet 228
Qy 738 -----GCTGAGGCCCTGCCCTGGGC-----CGG 761
Db 229 SerAsnPheSerLeuLeuLysGlyLeuHisLeuGlyAspThrGlyCysLeuLeuArgLys 248
Qy 762 GGAAGAGTTGGGCTGTGGAAGGACATCTTCACCGTCTCCATGAATGAAGCTTTGACTTG 821
Db 249 GlyThrProGlyAspTrpLysAsnTyrPheThrValAlaGlnAlaGluAlaPheAspLys 268

RESULT 74

US-11-152-635-10
; Sequence 10, Application US/11152635
; Publication No. US20060024708A1
; GENERAL INFORMATION:
; APPLICANT: Squires, James
; APPLICANT: Sinclair, Philip
; TITLE OF INVENTION: Porcine Sulfotransferase 2A1 Polynucleotide Sequence, Protein, and
; FILE OF INVENTION: Methods of Use for Same
; FILE REFERENCE: P06815US01
; CURRENT APPLICATION NUMBER: US/11/152,635
; CURRENT FILING DATE: 2005-06-14
; PRIOR APPLICATION NUMBER: 60/580,540
; PRIOR FILING DATE: 2005-06-17
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Porcine
US-11-152-635-10

Alignment Scores:
Pred. No.: 5.63e-20 Length: 285
Score: 347.50 Matches: 88
Percent Similarity: 50.3% Conservative: 56
Best Local Similarity: 30.8% Mismatches: 117
Query Match: 7.9% Indels: 25
DB: 6 Gaps: 6

US-10-768-158-1 (1-2419) x US-11-152-635-10 (1-285)

Qy 21 ATGGCGGAGAGCGAGCGCGAGACCCCGGAGTTCGAGACCAAGTACTTC 80
Db 1 MetThrGluGluGluVal-----ArgPheGluGlyIlePhePhe 13
Qy 81 GAGTTCATCGCTCGGCTCCGCCCTTCTGCGCGGGAAGATGGAGGATCGCCAA 140

Db 14 Pro---LysGlnIleLeuSerProGluMetLeuGlnGluValArgGluGlu----- 29
Qy 141 TTCCCGGTGGCGGCCAGCGAGCTGTGATCGTCACTTACCCCAAGTCCGCGACAGCTTG 200
Db 30 PheThrPheLysGluGluAspValLeuIleLeuThrPheProLysSerGlyThrAsnTrp 49
Qy 201 CTGCAGAGGTGTCTACTTGTGTGAGCCAGGCGCTGACCCCGATGAGATCGGCTTCATG 260
Db 50 MetIleGluIleIleCysLeuIleLeuSerLysGlyAspThrLysTrpIleGlnSerVal 69
Qy 261 AACATCAGCAGCAGCTCCCGCTCTCGGAGTACCCACAGCGCGGCTCGGACATCAAG 320
Db 70 ProAsnTrpAspArgSerProTrpLeuGlu---SerIleSerGlyTyrGluAsnLeuLys 88
Qy 321 GAACTGACCTCTCCCGCTCATCAAGAGCAGCTGCCCTACCGCTTCTGCTCTGAC 380
Db 89 AsnLysGluGlyProArgLeuIleSerSerHisLeuProIleGlnPheProLysAla 108
Qy 381 CTCCACATGGAGACTCCAAGTCTATATGCTCGCAACCCCAAGGATCTGCTGGTG 440
Db 109 PhePheLysSerLysAlaLysMetIleTyrIleileArgAsnProArgAspIleileVal 128
Qy 441 TCTTATTATCAGTTCACCGCTCTCTCGGACCATGAGTACCGAGGCACCTTTCAAGAA 500
Db 129 SerGlyTyrPhePheTrpLysSerThrAsnLeuValLysArgProGluSerLeuGluGln 148
Qy 501 TTCTGCCGGAGGTTTATGAATGATAAGCTGGGCTACCGCTCTCGTGTGTGACAGCTGCAG 560
Db 149 TyrPheGluTrpPheIleGlnGlyAsnValProTyrGlySerTrpPheAspHisIleArg 168
Qy 561 GAGTTCGGGAGCACCGCATGAGTCTCGAACGTGCTTTTCTCAAGTATGAGACATGCAT 620
Db 169 GlyTrpLeuProMetArgAspLysGluAsnValLeuIleLeuSerTyrGluGluLeuLys 188
Qy 621 CGGACCTGTGACGATGGTGAGCAGCTGCCAGATTCCTGGGGGTCTCCTGTGACAAG 680
Db 189 ArgAspThrArgSerAlaValGluLysIleCysGlnPheLeuGlyLysLysLeuGluPro 208
Qy 681 GCCCAGCTGGAAGCCCTGACGGAGCAGTGCACCGCTGGTGAGCAGTGTGCAAC--- 737
Db 209 GluGluLeuSerSerValValGluAsnSerSerPheGlnValMetLysGluAsnMet 228
Qy 738 -----GCTGAGGCCCTGCCCTGGGC-----CGG 761
Db 229 SerAsnPheSerLeuLeuLysGlyLeuHisLeuGlyAspThrGlyCysLeuLeuArgLys 248
Qy 762 GGAAGAGTTGGGCTGTGGAAGGACATCTTCACCGTCTCCATGAATGAAGCTTTGACTTG 821
Db 249 GlyThrProGlyAspTrpLysAsnTyrPheThrValAlaGlnAlaGluAlaPheAspLys 268

RESULT 75

US-09-795-926-2
; Sequence 2, Application US/09795926
; Patent No. US20020098486A1
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING THE SAME

QY 297 CAG-----CCGGCGCTGGACATC-----ATCAAGGAACGTGACC 329
Db 91 GluPheProValLeuGluCysGlyAspSerGluLysTyrGlnArgMetLysGlyPhePro 110
QY 330 TCTCCCGCCCTCATCAAGAGCCACCTGCGCTTCTGCTCCCTGACCTCCACCAAT 389
Db 111 SerProArgIleLeuAlaThrHisLeuHisTyrAspLysLeuProGlySerIlePheGlu 130
QY 390 GGAGACTCCAAGGTCTATATGCTCGCAACCCCAAGGATCTGGTGTCTTATTAT 449
Db 131 AsnLysAlaLysIleLeuValIlePheArgAsnProLysAspThrAlaValSerPheLeu 150
QY 450 CAGTTCCACCGCTCTCTCGGGACCACTGACCTACCGAGGACCTTTCAAGAATTTGCCGG 509
Db 151 HisPheHisAsnAspValProAspIleProSerTyrGlySerTyrAspGluPhePheArg 170
QY 510 AGGTTTATGAATGATAAGCTGGCTACGGCTCTCGTTTGAGCAGCTGCAGGAGTTCTGG 569
Db 171 GlnPheMetLysGlyGlnValSerTrpGlyArgTyrPheAspPheAlaIleAsnTrpAsn 190
QY 570 GAGCACCGCTGAGCTCGAACGTGCTTTTCTCAAGTATGAAGACATGATCGGACCTG 629
Db 191 LysHisLeuAspGlyAspAsnValLysPheIleLeuTyrGluAspLeuLysGluAsnLeu 210
QY 630 GTGACGATGCTGGAGCAGCTGCCAGATTCCTGCGGGGTCTCTGTGCAAGGCCACGCTG 689
Db 211 AlaAlaGlyIleLysGlnIleAlaGluPheLeuGlyPhePheLeuThrGlyGluGlnIle 230
QY 690 GAAGCCCTGACGGAGCACTGCCACAGCTGGTGGACCACTGCTGCAACGCTGAGGCCCTG 749
Db 231 GlnThrIleSerVal-----GlnSerThrPheGlnAlaMetArgAlaLysSerGln 247
QY 750 -----CCGCTGGGCG-----CGGGGAAGAGTTGGGCTGTGGGAG 782
Db 248 AspThrHisGlyAlaValGlyProPheLeuPheArgLysGlyGluValGlyAspTrpLys 267
QY 783 GACATCTTCACGCTCTCCATGAATGAGAGTTTGTGACTTGGTGTATPAAACAG 833
Db 268 AsnLeuPheSerGluIleGlnAsnGlnGluMetAspGluLysPheLysGlu 284

RESULT 77

US-10-072-012-664
; Sequence 664, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zernusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsbrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31

; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 664
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-072-012-664

Alignment Scores:
P. No.: 2,91e-16 Length: 303
Score: 303.50 Matches: 74
Percent Similarity: 48.7% Conservative: 61
Best Local Similarity: 26.7% Mismatches: 99
Query Match: 6.9% Indels: 43
DB: 4 Gaps: 9

US-10-768-158-1 (1-2419) x US-10-072-012-664 (1-303)
QY 75 TACTTCGAGTTCCATGCGTGGCGTCCG---CCCTTCTCGCGGGAAG---ATGGAG 128
Db 27 PhePheThrTyrGlnGlyIleProTyrProIleThrMetCysThrSerGluThrPheGln 46
QY 129 GAGATCGCCAACTTCCCGTGGCGGCCCGCCAGCAGCGTGGATCGTCACCTACCCCAAGTCC 188
Db 47 AlaLeuAspThrPheGluAlaArgHisAspAspIleValLeuAlaSerTyrProLysCys 66
QY 189 GGCACACAGC-----TTGCTGCGAGAGGTGGTCTACTTGTGTGAGCAGCGCGCT 236
Db 67 GlySerAsnTrpIleLeuHisIleValSerGluLeuIleTyrAlaValSerLys----- 84
QY 237 GACCCCGATGAGATCGGCTTGTATGAACATCGACGAGCAGCTCCCGGTCTCTGGAGTACCA 296
Db 85 -----LysLysTyrLysTyrPro 90
QY 297 CAG---CCGGCGCTGGACATC-----ATCAAGGAACGTGACC 329
Db 91 GluPheProValLeuGluCysGlyAspSerGluLysTyrGlnArgMetLysGlyPhePro 110
QY 330 TCTCCCGCCCTCATCAAGAGCCACCTGCGCTTCTGCTCCCTGACCTCCACCAAT 389
Db 111 SerProArgIleLeuAlaThrHisLeuHisTyrAspLysLeuProGlySerIlePheGlu 130
QY 390 GGAGACTCCAAGGTCTATATGCTCGCAACCCCAAGGATCTGGTGTCTTATTAT 449
Db 131 AsnLysAlaLysIleLeuValIlePheArgAsnProLysAspThrAlaValSerPheLeu 150
QY 450 CAGTTCCACCGCTCTCTCGGGACCACTGACCTACCGAGGACCTTTCAAGAATTTGCCGG 509
Db 151 HisPheHisAsnAspValProAspIleProSerTyrGlySerTyrAspGluPhePheArg 170
QY 510 AGGTTTATGAATGATAAGCTGGCTACGGCTCTCGTTTGAGCAGCTGCAGGAGTTCTGG 569
Db 171 GlnPheMetLysGlyGlnValSerTrpGlyArgTyrPheAspPheAlaIleAsnTrpAsn 190
QY 570 GAGCACCGCTGAGCTCGAACGTGCTTTTCTCAAGTATGAAGACATGATCGGACCTG 629
Db 191 LysHisLeuAspGlyAspAsnValLysPheIleLeuTyrGluAspLeuLysGluAsnLeu 210
QY 630 GTGACGATGCTGGAGCAGCTGCCAGATTCCTGCGGGGTCTCTGTGCAAGGCCACGCTG 689
Db 211 AlaAlaGlyIleLysGlnIleAlaGluPheLeuGlyPhePheLeuThrGlyGluGlnIle 230
QY 690 GAAGCCCTGACGGAGCACTGCCACAGCTGGTGGACCACTGCTGCAACGCTGAGGCCCTG 749
Db 231 GlnThrIleSerVal-----GlnSerThrPheGlnAlaMetArgAlaLysSerGln 247
QY 750 -----CCGCTGGGCG-----CGGGGAAGAGTTGGGCTGTGGGAG 782
Db 248 AspThrHisGlyAlaValGlyProPheLeuPheArgLysGlyGluValGlyAspTrpLys 267
QY 783 GACATCTTCACGCTCTCCATGAATGAGAGTTTGTGACTTGGTGTATPAAACAG 833
Db 268 AsnLeuPheSerGluIleGlnAsnGlnGluMetAspGluLysPheLysGlu 284

Db 171 GlnPheMetLysGlyGlnValSerTrpGlyArgTyrPheAspPheAlaIleAsnTrpAsn 190
Qy 570 GAGCACCAGCTGAGCTCGAACGTCGCTTTTCTCAAGTATGAAGACATGATCGGGACCTG 629
Db 191 LysHisLeuAspGlyAspAsnValLysPheIleLeuTyrGluAspLeuLysGluAsnLeu 210
Qy 630 GTGACGATGGTGGAGCAGCTGCCAGATTCTCTGGGGGTGCTCTGTGACAAGGCCAGCTG 689
Db 211 AlaAlaGlyIleLysGlnIleAlaGluPheLeuGlyPhePheLeuThrGlyGluGlnIle 230
Qy 690 GAAGCCCTGACGGAGCAGCTGCCACAGCTGTGGACACAGTCTGCAACGCTGAGGCCCTG 749
Db 231 GlnThrIleSerVal-----GlnSerThrPheGlnAlaMetArgAlaLysSerGln 247
Qy 750 -----CCCGTGGGC-----CGGGGAAGAGTTGGCTGTGAAG 782
Db 248 AspThrHisGlyAlaValGlyProPheLeuPheArgLysGlyGluValGlyAspTrpLys 267
Qy 783 GACATCTTCACCGTCTCCATGAATGAGAAGTTTGACTTGGTGTATAAACAG 833
Db 268 AsnLeuPheSerGluIleGlnAsnGlnGluMetAspGluLysPheLysGlu 284

RESULT 78

US-10-468-125-3
; Sequence 3, Application US/10468125
; Publication No. US20040082061A1
; GENERAL INFORMATION:
; APPLICANT: ASTROMOFF, Anna
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: DING, Li
; APPLICANT: DUGAN, Brendan M.
; APPLICANT: FORSYTHE, Ian J.
; APPLICANT: GIETZEN, Kimberly J.
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: LEE, Ernestine A.
; APPLICANT: LU, Yan
; APPLICANT: RICHARDSON, Thomas W.
; APPLICANT: RING, Huijun Z.
; APPLICANT: SANJANWALA, Madhusudan
; APPLICANT: SWARNAKAR, Anita
; APPLICANT: CHAWLA, Narinder K.
; APPLICANT: WARREN, Bridget A.
; APPLICANT: XU, Yuming
; APPLICANT: YUE, Henry
; APPLICANT: ZEBARJADIAN, Yeganeh
; TITLE OF INVENTION: DRUG METABOLIZING ENZYMES
; FILE REFERENCE: PI-0363 USN
; CURRENT APPLICATION NUMBER: US/10/468,125
; PRIOR FILING DATE: 2003-08-15
; PRIOR APPLICATION NUMBER: PCT/US02/04918
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/269,643
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/271,332
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: US 60/276,767
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: US 60/282,077
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: US 60/285,447
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: US 60/287,060
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/288,543
; PRIOR FILING DATE: 2001-05-03
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PERL Program
; SEQ ID NO 3
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:

; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 7491172CD1
US-10-468-125-3
Alignment Scores: 2.91e-16 Length: 303
Pred. No.: 303.50 Matches: 74
Score: 48.7% Conservative: 61
Best Local Similarity: 26.7% Mismatches: 99
Query Match: 6.9% Indels: 43
DB: 4 Gaps: 9
US-10-768-158-1 (1-2419) x US-10-468-125-3 (1-303)
Qy 75 TACTTCAGTTCATGCGGTGGCGTCCG---CCCTTCTCGCGCGGAAG---ATGAG 128
Db 27 PhePheThrTyrGlnGlyIleProTyrProIleThrMetCysThrSerGluThrPheGln 46
Qy 129 GAGATCGCAACTTCCCGGTGGCGGCCAGCGATGGATCGTCACCTACCCCAAGTCC 188
Db 47 AlaLeuAspThrPheGluAlaArgHisAspIleValLeuAlaSerTyrProLysCys 66
Qy 189 GGCACACAGC-----TTGCTCAGAGGTGCTCTACTTGGTGAGCCAGGGCGCT 236
Db 67 GlySerAsnTrpIleLeuHisIleValSerGluLeuIleTyrAlaValSerLys----- 84
Qy 237 GACCCGATGATCGGCTTGATGAACATCGACGAGCAGCTCCCGGTCTCTGGAGTACCA 296
Db 85 -----LysLysTyrLysTyrPro 90
Qy 297 CAG---CGGGCCTGGACATC-----ATCAAGGAATGACC 329
Db 91 GluPheProValLeuGluCysGlyAspSerGluLysTyrGlnArgMetLysGlyPhePro 110
Qy 330 TCTCCCGCTCATCAAGAGCACCTGCCCTTTCGCGCTTTCGCGCTCTGACCTCCACAT 389
Db 111 SerProArgIleLeuAlaThrHisLeuHisTyrAspLysLeuProGlySerIlePheGlu 130
Qy 390 GGAGACTCCAGGTCACTATATGGCTCGCAACCCCAAGGATCTGGTGTCTTATAT 449
Db 131 AsnLysAlaLysIleLeuValIlePheArgAsnProLysAspThrAlaValSerPheLeu 150
Qy 450 CAGTTCACCGCTCTCTCGCGACCATGAGCTACCGAGCACCTTTCAAGAAATTTGCGCG 509
Db 151 HisPheHisAsnAspValProAspIleProSerTyrGlySerTrpAspGluPhePheArg 170
Qy 510 AGGTTTATGAATGAAGCTGGGTACGGCTCTCGTTTGAACACGCTGACGAGTCTTGG 569
Db 171 GlnPheMetLysGlyGlnValSerTrpGlyArgTyrPheAspPheAlaIleAsnTrpAsn 190
Qy 570 GAGCACCAGCTGAGCTCGAACGTCGCTTTTCTCAAGTATGAAGACATGATCGGGACCTG 629
Db 191 LysHisLeuAspGlyAspAsnValLysPheIleLeuTyrGluAspLeuLysGluAsnLeu 210
Qy 630 GTGACGATGGTGGAGCAGCTGCCAGATTCTCTGGGGGTGCTCTGTGACAAGGCCAGCTG 689
Db 211 AlaAlaGlyIleLysGlnIleAlaGluPheLeuGlyPhePheLeuThrGlyGluGlnIle 230
Qy 690 GAAGCCCTGACGGAGCAGCTGCCACAGCTGTGGAGCCAGTCTGCAACGCTGAGGCCCTG 749
Db 231 GlnThrIleSerVal-----GlnSerThrPheGlnAlaMetArgAlaLysSerGln 247
Qy 750 -----CCCGTGGGC-----CGGGGAAGAGTTGGCTGTGAAG 782
Db 248 AspThrHisGlyAlaValGlyProPheLeuPheArgLysGlyGluValGlyAspTrpLys 267
Qy 783 GACATCTTCACCGTCTCCATGAATGAGAAGTTTGACTTGGTGTATAAACAG 833
Db 268 AsnLeuPheSerGluIleGlnAsnGlnGluMetAspGluLysPheLysGlu 284
RESULT 79
US-11-134-241-2
; Sequence 2, Application US/11134241

```
; Publication No. US20050287568A1
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Poter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; FILE REFERENCE: POLYNUCLEOTIDES ENCODING THE SAME
; CURRENT APPLICATION NUMBER: US/11/134,241
; PRIOR FILING DATE: 2005-05-20
; PRIOR APPLICATION NUMBER: US/10/364,774
; PRIOR FILING DATE: 2003-02-11
; PRIOR APPLICATION NUMBER: US/09/795,926
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 303
; TYPE: PRT
; ORGANISM: homo sapiens
; US-11-134-241-2

Alignment Scores:
Pred. No.: 2,91e-16 Length: 303
Score: 303.50 Matches: 74
Percent Similarity: 48.7% Conservative: 61
Best Local Similarity: 26.7% Mismatches: 99
Query Match: 6.9% Indels: 43
DB: 6 Gaps: 9

US-10-768-158-1 (1-2419) x US-11-134-241-2 (1-303)
QY 75 TACTTCGAGTTCATGGGTGGCGTCCG---CCCTTCTCGCGGGAAG---ATGAG 128
Db 27 PhePheThrTyrGlnGlyIleProTyrProIleThrMetCysThrSerGluThrPheGln 46
QY 129 GAGATCGCAACTTCCCGTCCGCGCCAGCGATGTCGTCACCTACCCCAAGTCC 188
Db 47 AlaleuAspThrPheGluAlaArgHisAspAspIleValLeuAlaSerTyrProLysCys 66
QY 189 GGCACCCAGC-----TTGCTCAGAGGTGGTCTACTTGGTGAGCCAGGGCGCT 236
Db 67 GlySerAsnTrpIleLeuHisIleValSerGluLeuIleTyrAlaValSerLys----- 84
QY 237 GACCCCGATGAGTCGGCTTGATGAACATCGACGAGCGCTCCCGGTCTCGAGTACCCA 296
Db 85 -----LysLysTyrLysTyrPro 90
QY 297 CAG---CCGGCCCTGGACATC-----ATCAAGGAACGTACC 329
Db 91 GluPheProValLeuGluCysGlyAspSerGluLysTyrGlnArgMetLysGlyPhePro 110
QY 330 TCTCCCGCCTCATCAAGAGCCACTGCCCTACCGCTTCTGCGCTCTGACCTCCACAA 389
Db 111 SerProArgIleLeuAlaThrHisLeuHisTyrAspLysLeuProGlySerIlePheGlu 130
QY 390 GGAGATCCAGGTCACTATATGGCTCGCAACCCCAAGGATCTGGTGGTCTCTATTAT 449
Db 449 -----LysLysTyrLysTyrPro 90
```

```
Db 131 AsnLysAlaLysIleLeuValIlePheArgAsnProLysAspThrAlaValSerPheLeu 150
QY 450 CAGTTCACCGCTCTCTCGCGGACCATGAGTACCGAGGACACCTTTCAAGAATTTCGCG 509
Db 151 HisPheHisAsnAspValProAspIleProSerTyrGlySerTyrAspGluPhePheArg 170
QY 510 AGTTTATGATGATAAGCTGGCTACGGCTCCCTGGTTTGAGCAGCTGCAGGATTTCTGG 569
Db 171 GlnPheMetLysGlyGlnValSerTrpGlyArgTyrPheAspPheAlaIleAsnTrpAsn 190
QY 570 GAGCACCGCATGGACTCGAAGCTCTTCTCAAGTATGAAGACATGCGATCGGGACCTG 629
Db 191 LysHisLeuAspGlyAspAsnValLysPheIleLeuTyrGluAspLeuLysGluAsnLeu 210
QY 630 GTGACGATGGTGAGCAGCTGCGCAGATTCCTGGGGGTGTCTCTGTGCAAGGCCAGCTG 689
Db 211 AlaAlaGlyIleLysGlnIleAlaGluPheLeuGlyPheLeuThrGlyGluGlnIle 230
QY 690 GAAGCCCTGACGGAGCACTGCCACCACTGGTGACAGCTGCTGCAAGCGCTGAGGCCCTG 749
Db 231 GlnThrIleSerVal-----GlnSerThrPheGlnAlaMetArgAlaLysSerGln 247
QY 750 -----CCCGTGGGC-----CGGGGAAGAGTTGGGCTGTGGAG 782
Db 248 AspThrHisGlyAlaValGlyProPheLeuPheArgLysGlyValGlyAspTrpLys 267
QY 783 GACATCTTCAACGCTCTCCATGAATGAGAAGTTTCACTTGGTGTATAACAG 833
Db 268 AsnLeuPheSerGluIleGlnAsnGlnGluMetAspGluLysPheLysGlu 284

RESULT 80
US-10-092-900A-150
; Sequence 150, Application US/10092900A
; Publication No. US20040043382A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Taupier Jr., Raymond J.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Ji, Weizhen
; APPLICANT: Gorman, Linda
; APPLICANT: Miller, Charles E.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Guo, Xiaojia Sashu
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Fernandes, Elma R.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Liu, Yi
; APPLICANT: Anderson, David W.
; APPLICANT: Spaderna, Steven K.
; APPLICANT: Catterton, Elina
; APPLICANT: Leite, Mario W.
; APPLICANT: Zhong, Haihong
; APPLICANT: Alsobrook, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: No. US20040043382A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-290C
; CURRENT APPLICATION NUMBER: US/10/092,900A
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: USN 60/274,322
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USN 60/283,675
```

```

; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: USN 60/338,092
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: USN 60/274,281
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USN 60/274,191
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USN 60/325,681
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: USN 60/304,354
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: USN 60/279,995
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: USN 60/294,899
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: USN 60/287,424
; PRIOR FILING DATE: 2001-04-30
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 768
; SEQ ID NO 150
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-900A-150

Alignment Scores:
Pred. No.: 4,67e-16 Length: 295
Score: 301.00 Matches: 85
Percent Similarity: 47.8% Conservative: 58
Best Local Similarity: 28.4% Mismatches: 120
Query Match: 6.8% Indels: 36
DB: 4 Gaps: 8

US-10-768-158-1 (1-2419) x US-10-092-900A-150 (1-295)
Qy 30 AGCGAGCGGAGCCGCCAGCAGCCCGGGGAGTTCGAGAGCAAGTACTTCGAGTTCAT 89
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; 7 SerGluLeuGluTrpGlyLeuLeuProGlu-----GluPheSerGlnValasn 23
Qy 90 GCGCTGCTGCTGCGCCCTCTCGCGCGGAGAGATGAGAGATCGCAACTTCCCGGTG 149
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; 24 GlyIleLeuGlnLysLysMetCysAspPheTrpAspLysIleTrpAsnPhaGlnAla 43
Qy 150 CGGCCAGCGAGTGTGATCGTACCTACCCAGTCCGCGAGTCCGAGTTCCTCCAGGAG 209
Db :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
; 44 LysProAspAspLeuLeuIleAlaSerTyrProLysAlaGlyThrTrpThrGlnGlu 63
Qy 210 GTGCTCTACTGTGAGCGCGCGCTGACCCCGATGAGATCGGCTTGATGAACATCGAC 269
Db :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
; 64 IleValAspLeuIleGlnAsnAspGlyAspIleGluLysSerArgAlaSerIleGln 83
Qy 270 GAGCAGCTCCCGTCTCGAGTACCCACAG-----CCGGGC 305
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; 84 LeuGlnHisProPheLeuGluTrpIleArgMetThrHisAlaArgLysIlePheAlaGly 103
Qy 306 CTGACATCATCAGGAACCTACCTCCCGCCTCATCAGAGCCACCTGCCCTACCGC 365
Db :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
; 104 IleAspGlnAlaAsnThrMetProSerProArgThrLeuLysThrHisLeuProValGln 123
Qy 366 TTCTGCGCTCTGACCTCCACAATGGAGCTCCAAAGGTCTATATGCTCGCAACCCC 425
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; 124 LeuLeuProProSerPheTrpGluGlnAsnCysLysIleIleTyrValAlaArgAsnAla 143
Qy 426 AAGGATCTGGTGTCTTATTATCAGTCCACCGCTCTCTCGGACCATGAGTACCGA 485
Db ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
; 144 LysAspAsnLeuValSerTyrTyrHisPheGlnArgMetSerLysAlaLeuProAspVal 163
Qy 486 GGCACCTTCAAGAACTTCGCGGAGTTTATGATGATAGCTGGGTACGCTCCTCGG 545
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; 164 LeuThrValGlyGluTyr-----IleMetCysGlyGluValLeuTrpGlyIleTrp 180
Qy 546 TTTGAGCAGCTGCAGGAGTTCTGGGAGCAGCCGCGATGGAGCTGCTCTTTCTCAAG 605
Db ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

US-10-768-158-1 (1-2419) x US-11-097-143-11370 (1-313)
Qy 132 ATGCCCAACTTCCCGTGGCGGCCCGAGCGTGTGGATCGTACCTACCCCAAGTCCGGC 191
Db :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
; 38 ValHisAspMetLysLeuArgAspAspValTrpIleValThrLeuProLysCysGly 57
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

US-10-768-158-1 (1-2419) x US-11-097-143-11370 (1-313)
Pred. No.: 1,53e-15 Length: 313
Score: 295.00 Matches: 83
Percent Similarity: 48.7% Conservative: 51
Best Local Similarity: 30.2% Mismatches: 103
Query Match: 6.7% Indels: 38
DB: 6 Gaps: 8

```



```
Qy 780 AAGGACATCTCACCGTCTCCATGAATGAGAAGTTTGACTTGGTGTATAAACAG 833
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 229 LysAsnLeuPheSerGluIleGlnAsnGlnGluMetAspGluLysPheLysGlu 246
US-10-364-774-6
; Sequence 6, Application US/10364774
; Publication No. US20030144497A1
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; FILE REFERENCE: LEX-0144-USA
; CURRENT APPLICATION NUMBER: US/10/364,774
; PRIOR FILING DATE: 2003-02-11
; PRIOR APPLICATION NUMBER: US/09/795,926
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 265
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-364-774-6
Alignment Scores:
Pred. No.: 2,33e-15 Length: 265
Score: 292.50 Matches: 70
Percent Similarity: 49.2% Conservative: 57
Best Local Similarity: 27.1% Mismatches: 90
Query Match: 6.6% Indels: 41
DB: 4 Gaps: 7
US-10-768-158-1 (1-2419) x US-10-364-774-6 (1-265)
Qy 126 GAGGATCGCAACTCCCGTGGCGCCAGCGACGTGTGATCGTCACCTACCCCAAG 185
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 8 GlnAlaLeuAspThrPheGluAlaArgHisAspAspIleValLeuAlaSerTyProls 27
Qy 186 TCCGCACACAGC-----TTGCTGCAGAGGTGGTCTACTTGTGTAGCAGCGGC 233
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 28 CysGlySerAsnTrpIleLeuHisIleValSerGluLeuIleTyAlaValSerLys--- 46
Qy 234 GCTGACCCCGATGAGATCGGCTTGATGAACATCGACGACGCTCCCGCTCTGGAGTAC 293
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 47 -----LysLysTyLysTyLys 51
Qy 294 CCACAG---CCGGGCTCGACATC-----ATCAAGGAACGTG 326
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 52 ProGluPheProValLeuGluCysGlyAspSerGluLysTyArgMetLysGlyPhe 71
Qy 327 ACCTCTCCCGGCTCATCAAGACACCTGCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTC 386
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 72 ProSerProArgIleLeuAlaThrHisLeuHisTyAspLysLeuProGlySerIlePhe 91
Qy 387 AATGAGAGACTCCAAAGTTCATCTATATGCTCGCAACCCCAAGGATCTGTGTGTCTTAT 446
```

```
Db 92 GluAsnLysAlaLysIleLeuValIlePheArgAsnProLysAspThrAlaValSerPhe 111
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 447 TATCAGTTCCACCGCTCTCTGCGGACCATGAGCTACCGAGCACCTTTCAAGATTTCGC 506
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 112 LeuHisPheHisAsnAspValProAspIleProSerTyArgTyPheAspGluPhePhe 131
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 507 CGGAGGTTTATGATGATAAGCTGGCTACGGCTCCCTGGTTTGAGCACGTCGACGAGTTC 566
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 132 ArgGlnPheMetLysGlyGlnValSerTrpGlyArgTyPheAspPheAlaIleAsnTrp 151
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 567 TGGAGCACCGCATGGACTCGAACGCTGTTTCTCAAGTATGAAGCATGCATCGGGAC 626
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 152 AsnLysHisLeuAspGlyAspAsnValLysPheIleLeuTyGluAspLeuLysGluAsn 171
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 627 CTGCTGACGATGGTGGAGCAGCTGGCCAGATTCTCTGGGGGTGCTCTGTGACAAAGGCCAG 686
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 172 LeuAlaAlaGlyIleLysGlnIleAlaGluPheLeuGlyPhePheLeuThrGlyGluGln 191
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 687 CTGGAAGCCCTGACGGAGCACTGCCACAGCTGGTGGACAGCTCTCAACGCTGAGGCC 746
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 192 IleGlnThrIleSerVal-----GlnSerThrPheGlnAlaMetArgAlaLysSer 208
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 747 CTG-----CCCGTGGGC-----CGGGAAGAGTTGGGCTGTGG 779
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 209 GlnAspThrHisGlyAlaValGlyProPheLeuPheArgLysGlyGluValGlyAspTrp 228
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 780 AAGGACATCTTCACCGCTCTCCATGAATGAGAAGTTTGACTTGGTGTATAAACAG 833
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 229 LysAsnLeuPheSerGluIleGlnAsnGlnGluMetAspGluLysPheLysGlu 246
RESULT 84
US-11-134-241-6
; Sequence 6, Application US/11134241
; Publication No. US20050287568A1
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; FILE REFERENCE: LEX-0144-USA
; CURRENT APPLICATION NUMBER: US/11/134,241
; PRIOR FILING DATE: 2005-05-20
; PRIOR APPLICATION NUMBER: US/10/364,774
; PRIOR FILING DATE: 2003-02-11
; PRIOR APPLICATION NUMBER: US/09/795,926
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 265
; TYPE: PRT
; ORGANISM: homo sapiens
US-11-134-241-6
Alignment Scores:
Pred. No.: 2.33e-15 Length: 265
```

Score:	232.50	Matches:	70
Percent Similarity:	49.2%	Conservative:	57
Best Local Similarity:	27.1%	Mismatches:	90
Query Match:	6.6%	Indels:	41
DB:	6	Gaps:	7

US-10-768-158-1 (1-2419) x US-11-134-241-6 (1-265)			
QY	126	GAGGAGATCGCAACATTC	CGCGGTGGCGCCACGACGCGTGTGATCGTCACTACCCCAAG 185
		::: :::	: ::: :::
Db	8	GlnAlaLeuAspThrPheGluAlaArgHisAspAlleValLeuAlaSerTyrProLys 27	
QY	186	TCGCGCACACAGC-----	TTGCTGCAGGAGGTGTCTACTTGTGTGACCCAGGCC 233
		: :::	: :::
Db	28	CysGlySerAsnTrpIleLeuHisIleValSerGluLeuIleTyrAlaValSerLys--- 46	
QY	234	GCTGACCCCGATGAGATCGGCTTGATGAACATCGACGACGCTCCGCGCTCGGAGTAC 293	
		: :::	: :::
Db	47	-----	-----LysLysTyrLysTyr 51
QY	294	CCACAG---CCGGGCTTGGACATC-----	ATCAAGGAACTG 326
		: :::	: :::
Db	52	ProGluPheProValLeuGluCysGlyAspSerGluLysTyrGlnArgMetLysGlyPhe 71	
QY	327	ACCTCTCCCGGCTCATCAGAGCACCTCGCCTTACCGCTTTCGCGCTCTGCACCTCCAC 386	
		: :::	: :::
Db	72	ProSerProArgIleLeuAlaThrHisLeuHisTyrAspLysLeuProGlySerIlePhe 91	
QY	387	AATGGAGATCCCAAGTCACTATATGCTCGCAACCCCAAGGATCTGGTGGTCTTAT 446	
		: :::	: :::
Db	92	GluAsnLysAlaLysIleLeuValIlePheArgAsnProLysAspThrAlaValSerPhe 111	
QY	447	TATCAGTTCCACCGCTCTCTCGCGACCATGAGTACCGAGGCACCTTTTCAAGAAATTTCTGC 506	
		: :::	: :::
Db	112	LeuHisPheHisAsnAspValProAspIleProSerTyrGlySerTyrAspGluPhePhe 131	
QY	507	CGGAGGTTTATGAATGAATGCTGGGCTACGGCTCTCGTGTGTGACAGCTGCAGGAGTTC 566	
		: :::	: :::
Db	132	ArgGlnPheMetLysGlyGlnValSerTrpGlyArgTyrPheAspPheAlaIleAsnTrp 151	
QY	567	TGGGACACCGCATGGACTCGCAACGTGCTTTTCTCAAGTATGAGACATGCATCGGGAC 626	
		: :::	: :::
Db	152	AsnLysHisLeuAspGlyAspAsnValLysPheIleLeuTyrGluAspLeuLysGluAsn 171	
QY	627	CTGGTCACGATCGTGAGCAGCTGGCCAGATTCTTGGGGGTGCTCTGTGCAAGGCCAC 686	
		: :::	: :::
Db	172	LeuAlaAlaGlyIleLysGlnIleAlaGluPheLeuGlyPhePheLeuThrGlyGluGln 191	
QY	687	CTGGAAGCCCTGACGGAGCATGCGCACCGACTGGTGGACCAAGTCTGCCACGCTGAGGCC 746	
		: :::	: :::
Db	192	IleGlnThrIleSerVal-----	-----GlnSerThrPheGlnAlaMetArgAlaLysSer 208
QY	747	CTG-----CCCGTGGGC-----	-----CGGGCAAGAGTTGGGCTGTGG 779
		: :::	: :::
Db	209	GlnAspThrHisGlyAlaValGlyProPheLeuPheArgLysGlyGluValGlyAspTrp 228	
QY	780	AAGGACATCTTCACCGTCTCCATGAATCAGAAAGTTTGACTTGGTGATAAACAG 833	
		: :::	: :::
Db	229	LysAsnLysPheSerGluIleGlnAsnGlnGluMetAspGluLysPheLysGlu 246	

RESULT 85
US-10-072-012-665
; Sequence 665, Application US/10072012
; Publication No. US2004003493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimketa, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangoli, Esha
; APPLICANT: Padigar, Muralidhara

```
QY 306 -----CTGGACATCATCAAGAACTGACCTCTCCCGCCTCATCAAGACCACTG 356
Db 109 ThrGluLysTyrGluArgMetLysLysLeuProSerArgArgValIleLeuThrHisLeu 128
QY 357 CCTACCGCTTCTGCGCCTCTGACCTCCACAATGGAGACTCCAAGGTCTATATGGCT 416
Db 129 SerProGluLysLeuProLysSerIlePheLysAsnLysAlaLysIleLeuLeuLeu 148
QY 417 CGCAACCCCAAGATCTGGTGGTCTTATTATCATGTTCCACCGCTCTCTCGGACCATG 476
Db 149 ArgAsnProLysAspIleAlaThrSerPhePheHisPheSerAsnArgTrpSerAlaLeu 168
QY 477 AGCTACCGAGCACTTTCAGAAATTCGCGGAGGTTTATGAATGATAAGCTGGGTAC 536
Db 169 ProSerTyrGluThrTrpAspAspPhePheIleAlaPheMetThrGluLysMetProTrp 188
QY 537 GGCTCTCTGTTGAGCAGCTGAGAGTCTCGGAGACACCGCATGGACTCGAAGTGTCT 596
Db 189 GlySerTyrPheAsnTyrLeuSerGluTrpAsnLysTyrAlaAlaAspGluAsnValMet 208
QY 597 TTTCTCAAGTAAAGACATCATCGGACCTGGAGCTGTGACGATGGTGAGCAGCTGGCCAGA 656
Db 209 ThrIleThrTyrGluGluLysGluAsnGlnThrLeuGlyValLysAsnIleAlaSer 228
QY 657 TTCTGGGGGTCTCTGTGACAGCCCGCAGCTGGAGCCCTGACGGGACACTGC----- 710
Db 229 PhePheGlyIleSerLeuThrGlyGluGluLeuArgSerValIleGluArgSerSerPhe 248
QY 711 CACCACTGGTGGACCACTGTGCAACGCTGAG-----GCCCTGCCGTGGGC 758
Db 249 GlnSerMetLysGluAsnSerLeuLysThrHisGlyAlaLeuGlySerMetLeuPheArg 268
QY 759 CGGGAAGAGTTGGCTGTGGAAGACATCTTACCGCTCTCCATGATGAGAAGTTTGAC 818
Db 269 LysGlyGlyValSerAspTrpLysAsnLeuPheAsnGluGluGlnAsnGluLysMetAsp 288
QY 819 TTGCTGTATAACAGAAATGGGAAGTGTGACCTC 854
Db 289 LysValPheGluGluArgIleAlaArgThrLysLeu 300

RESULT 86
US-10-072-012-250
; Sequence 250 Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shinkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31
```

```
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 250
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-072-012-250

Alignment Scores:
Pred. No.: 2,06e-14 Length: 303
Score: 281.50 Matches: 72
Percent Similarity: 48.0% Conservative: 61
Best Local Similarity: 26.0% Mismatches: 101
Query Match: 6.4% Indels: 43
DB: 4 Gaps: 9

US-10-768-158-1 (1-2419) x US-10-072-012-250 (1-303)
QY 75 TACTTCGAGTTCCATGCGTGGCTCCG---CCCTTCTGCCCGGGAG---ATGGAG 128
Db 27 PhePheThrTyrGlnGlyIleProTyrProIleThrMetCysThrSerGluThrPheGln 46
QY 129 GAGATCGCAACTTCCCGTGGCGCCAGCGAGCTGGATCGTCACCTACCCCAAGTCC 188
Db 47 AlaLeuAspThrPheGluAlaArgHisAspIleValLeuAlaSerTyrProLysCys 66
QY 189 GGCACACAGC-----TTGCTCAGAGAGTGGTCTACTTTGGTGAGCCAGCGGCT 236
Db 67 GlySerAsnTrpIleLeuHisIleValSerGluLeuIleTyrAlaValSerLys----- 84
QY 237 GACCCCGATGAGTCGGCTTGATGAACATCGACGAGCAGCTCCCGGCTCTGGAGTACCCA 296
Db 85 -----LysLysTyrLysTyrPro 90
QY 297 CAG---CGGGCCTGGACATC-----ATCAAGGAACCTGACC 329
Db 91 GluPheProValLeuGluCysGlyAspSerGluLysTyrGlnArgMetLysGlyPhePro 110
QY 330 TCTCCCGCCTCATCAAGAGCCACCTGCCCTTACCGCTTTCTGCGCTCTGACCTCCACAA 389
Db 111 SerProArgIleLeuAlaThrHisLeuHisTyrAspLysLeuProGlySerIlePheGlu 130
QY 390 GGAGACTCCAAGTCTATATATGCTCGCAACCCCAAGGATCTGGTGGTCTTATTAT 449
Db 131 AsnLysAlaLysIleLeuValIlePheArgAsnProLysAspThrAlaValSerPheLeu 150
QY 450 CAGTTCCACCGCTCTCTCGGAGCACCATGAGCTACCGAGGCACCTTTCAAGAAATTCGCGG 509
Db 151 HisPheHisAsnAspValProAspIleProSerTyrGlySerTyrAspGluPhePheArg 170
QY 510 AGGTTTATGAATGATAAGCTGGGTACGGCTCTCGGTTTGAGCACGTCGACGAGGTTCTGG 569
Db 510 AGGTTTATGAATGATAAGCTGGGTACGGCTCTCGGTTTGAGCACGTCGACGAGGTTCTGG 569
```



```
i OTHER INFORMATION: Ceres Seq. ID no. 15218612
US-11-096-568A-1016

Alignment Scores:
Pred. No.: 5,18e-14 Length: 345
Score: 277.00 Matches: 84
Percent Similarity: 42.8% Conservative: 43
Best Local Similarity: 28.3% Mismatches: 43
Query Match: 6.3% Indels: 43
DB: 6 Gaps: 10

US-10-768-158-1 (1-2419) x US-11-096-568A-1016 (1-345)
QY 53 CCGCGGGAGTTCGAGACCAAGTACTTCGAGTTCCA-----TGGCGTGGCGTCCGCC 106
Db 39 ProSerGlyThrLeuGlnAspTyrCysArgHisProSerGlyTrpCysIleThrLeu-Pr 58
QY 107 CTTCTCCCGCGGAAGATGGAGGAGATCGCAACTTCCCGTGGCGCCAGCGAGCTGTG 166
Db 58 oileMetValSerSerMetValAlaGluGlnHisPheGluAlaArgGlyThrAspValLe 78
QY 167 GATCGTCACTACCCCAAGTCGGCACCGAGCTTGTGCGAGGAGTGTCTACTTGGTGAG 226
Db 78 uLeuValThrMetProLysSerGlyThrThrTrpIleLysAlaLeuLeuTyAlaAlaAl 98
QY 227 CCAGGCGCTGACCCCGATGAGATCGGCTTGATG-----AACATCGA 268
Db 98 aHisArgThrAspAspThrSerSerSerIleLeuArgGlnLeuAlaSerHisAsnSerHi 118
QY 269 CGAGCAGCTCCCGTCTCGGAG-----TACCACAGCGCGGCTGGACATCATCAA 319
Db 118 sGlnLeuValProPheLeuGluAlaGlnValTyrThrLysAspGlnIleProAspLeuSe 138
QY 320 GAACTGACCTCTCCCGGCTCATCAAGAGCACCTGCGCTTACCGCTTCTGCGCTCTGA 379
Db 138 rSerLeuProAlaProArgLeuPheAlaThrHisIleProAlaGluSerLeuProProSe 158
QY 380 CTTCCACATGGAGACTCCAGGCTCATCTATATGCTCGCAACCCCAAGGATCTGGTGGT 439
Db 158 rValValAlaSerGlyCysLysValValTyrLeuCysArgAspProLysAspCysPheVa 178
QY 440 GTCTATTATCATGTTCCACCGCTCTCGGACCATGAGCTACCGAGGACCCCTTTCAAGA 499
Db 178 lSerLeuTrpHisPheMetAsnLysPheThrProTrpAspIleAspGluAlaHisGlyAr 198
QY 500 ATTCTCCGGAGGTTTATGAATGATGAAGCTGGGC-----TACGGCTCTCGTTTGA 550
Db 198 gPheCysGlu-----GlyValSerLeuTyrGlyProPheTrpGl 211
QY 551 GCACGTGAGGAGTCTGGGACACCGCATGGAC-----TCGAACGTCTTTTCTCAA 604
Db 211 uHisValLeuSerTyrTrpArgTrpHisValAspArgProGlyGlnValLeuPheLeuTh 231
QY 605 GTATCAAGACATGATCGGACCTCGTGACGATGGTGAGCAGCTGGGACAGATTCTCTGG 664
Db 231 rTyrGluGluLeuSerAlaAspProLeuGlyGlnLeuArgLeuAlaGluPheIleGl 251
QY 665 GGTGCTCTGTGAAGGCCAGCTGGAGCC-----CTGACGGACACTG 709
Db 251 yArgProPheThrProGlyGluGlnGluAlaGlyValAspArgGluIleAlaGluAlaCy 271
QY 710 C-----CACCAGCTGGTGACCAAGTCTGCTCAACGCTGAG-----743
Db 271 salAMetLysSerMetValAsnGlnGluValAsnGlnSerArgThrThrGluIleValGl 291
QY 744 ----GCCCTCGCGTGGG-----CGGGAGAGATTGGGCTGTGTGAAGACAT 787
Db 291 uMetProIleProAsnGlyIlePhePheArgArgGlyValValGlyAspTrpThrAsnTy 311
QY 788 CTTCCACGGTCTCCATGAATGAGAAGTTGACTTGGTGTGTATTAACAGAG 836
Db 311 rLeuThrProGluMetAlaGlyArgIleAspGluIleThrLysSerLys 327
```

```
RESULT 89
US-09-854-122-20
; Sequence 20 Application US/09854122
; Patent No. US20020016980A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTE, RANDALL S.
; APPLICANT: SMITH, ROBERT
; TITLE OF INVENTION: TRANSGENIC PLANTS INCORPORATING TRAIT OF ZOSTERA MARINA
; FILE REFERENCE: PHA-007.01
; CURRENT APPLICATION NUMBER: US/09/854,122
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/202,529
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Flaveria bidentis
US-09-854-122-20

Alignment Scores:
Pred. No.: 6.74e-14 Length: 320
Score: 275.50 Matches: 80
Percent Similarity: 49.1% Conservative: 63
Best Local Similarity: 27.5% Mismatches: 111
Query Match: 6.3% Indels: 37
DB: 3 Gaps: 11

US-10-768-158-1 (1-2419) x US-09-854-122-20 (1-320)
QY 72 AAGTACTTCGAGTTCCATGCGGTGCGGCTTCTGCGCGGAGAGATGAGGAG 131
Db 34 LysTyrGlnAspPheTrpGly-----LeuGlnAsnAsnIleGluGlyAlaIleLeuAla 51
QY 132 ATCCCAACTTCCCGGTGCGGCCCGAGCGAGCTGTGGATGTCACCTACCCCAAGTCCGGC 191
Db 52 GlnGlnSerPheLysAlaArgProAspAspValPheLeuCysSerTyrProLysSerGly 71
QY 192 ACCAGCTGTGCGAGGAGTGTCTACTTGGTGGAGCAGGCGGTGAGCCCGATGATC 251
Db 72 ThrTrpLeuLysAlaLeuAlaTyrAlaIleValThrArgGluLysPheAspGluPhe 91
QY 252 GGC-----TTGATGAACATCGACGAGCAG-----CTCCCGGTCTGGAGTACCCACAG 299
Db 92 ThrSerProLeuLeuThrAsnIleProHisAsnCysIleProTyrIleGlu-----108
QY 300 CCGGCGCTTGGACATCAAGGAACTGACCTCTCCCCGC-----CTCATC 344
Db 109 -----LysAspLeuLysLysIleValGluAsnGlnAsnSerCysPheThrProMet 126
QY 345 AAGAGCCACCTGCGCTTCCCGCTTCTGCGCTTCTGACCTCCACATGGAGACTCCAGGTC 404
Db 127 AlaThrHisMetProTyrHisValLeuProLysSerIleLeuAlaLeuAsnCysLysMet 146
QY 405 ATCTATATGGCTCGCAACCCCAAGGATCTGTGTGTCTTATTATTCAGTTCACCGCTCT 464
Db 147 ValTyrIleTyrArgAsnIleLysAspValIleValSerPheTyrHisPheGlyArgGlu 166
QY 465 CTGCGGACCATGAGTACCGA---GGCACCTTTCAAGAAATTCCTCGCGAGGTTTATGAAT 521
Db 167 IleThrLysLeuProLeuGluAspAlaProPheGluAlaPheAspGluPheTyrHis 186
QY 522 GATTAAGCTGGGCTACGGCTCTCTGTTTGGACGACCTGCGAGGAGTCTGGGAGCACCGCATG 581
Db 187 GlyIleSerGlnPheGlyProTyrTrpAspHisLeuLeuGlyTyrTrpLysAlaSerLeu 206
QY 582 GAC-----TCGAACGTCTTTTCTCAAGTATGAAGACATGCATCGCGGACCTGTGTGACG 635
Db 207 GluArgProGluValIleLeuPheLeuLysTyrGluAspValLysLysAspProThrSer 226
QY 636 ATGTGTGAGCAGCTGGCCAGATTCTCTGGGGGTGTCTGT-----GACAAG 680
```

Db 227 AsnVallysArgLeuAlaGluPheIleGlyTyrProPheThrPheGluGluGluLysGlu 246
Qy 691 GCCAGCTGGAGCCCTGACGGACACTGC-----CACCAGCTGGTG 722
Db 247 GlyValIleGluSerIleIleLysLeuCySerPheGluAsnLeuSerAsnLeuGluVal 266
Qy 723 GACCACTGCTGCAACCTGAGGCC---CTGCCCGTG-----GGCCGG 761
Db 267 AsnLysSerGlyAsnSerLysGlyPheLeuProIleGluAsnArgLeuTyrPheArgLys 286
Qy 762 GGAAGAGTTGGCTGTGAAGACATCTTACCCTCTCATGAATGAGAAGTTTGCATTG 821
Db 287 AlalysAspGlyAspTrpLysAsnTyrPheThrAspGluMetThrGluLysIleAspLys 306
Qy 822 GTGTATAACAGAAAGATCGGAAGTGTGACCTC 854
Db 307 LeuIleAspGluLysLeuSerAlaThrGlyLeu 317
RESULT 90
US-11-033-030-20
; Sequence 20, Application US/11033030
; Publication No. US20060053510A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTE, RANDALL S.
; APPLICANT: SMITH, ROBERT
; TITLE OF INVENTION: TRANSGENIC PLANTS INCORPORATING TRAIT OF ZOSTERA MARINA
; FILE REFERENCE: PHA-007.01
; CURRENT APPLICATION NUMBER: US/11/033,030
; PRIOR FILING DATE: 2005-01-11
; PRIOR APPLICATION NUMBER: US/09/854,122
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/202,529
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Flaveria bidentis
US-11-033-030-20
Alignment Scores:
Pred. No.: 6.74e-14 Length: 320
Score: 275.50 Matches: 80
Percent Similarity: 49.1% Conservative: 63
Best Local Similarity: 27.5% Mismatches: 111
Query Match: 6.3% Indels: 37
DB: 6 Gaps: 11
US-10-768-158-1 (1-2419) x US-11-033-030-20 (1-320)
Qy 72 AAGTACTTCGAGTTCCATGGCGTGGCTGCGCCCTTCTGCGCGGAAGATGAGGAG 131
Db 34 LysTyrGlnAspPheTrpGly-----LeuGlnAsnAsnIleGluGlyAlaIleLeuAla 51
Qy 132 ATCGCCAACTTCCGGTGGCGGCCAGCGAGCGATGATCGTACCTACCCCAAGTCGGC 191
Db 52 GlnGlnSerPheLysAlaAspProAspAspValPheLeuCySerTyrProLysSerGly 71
Qy 192 ACCAGCTTGTGCGAGGAGTGTCTACTTGTGTGAGCCAGGCGCTGACCCCGATGAGATC 251
Db 72 ThrThrTrpLeuLysAlaLeuAlaTyrAlaIleValThrArgGluLysPheAspGluPhe 91
Qy 252 GGC-----TTGATGAACATCAGCAGGAGCAG---CTCCCGTCTCTGGAGTACCCACAG 299
Db 92 ThrSerProLeuLeuThrAsnIleProHisAsnCySllieProTyrIleGlu----- 108
Qy 300 CCGGCGCTGGACATCATCAAGAACTGACCTCTCCCCGC-----CTCATC 344
Db 109 -----LysAspLeuLysIleValGluAsnGlnAsnSerCySllieProMet 126
Qy 345 AAGAGCCACCTGCGCTTCTGCGCTTCTGACCTTCCCAATGGAGACTCCAAAGTC 404
DB: 111 Indels: 111 Gaps: 111

Db 127 AlaThrHisMetProTyrHisValLeuProLysSerIleLeuAlaLeuAsnCysLysMet 146
Qy 405 ATCTATATGGCTGCAACCCCAAGGATCTGTGTGTCTTATTATTCAGTTCACCGCTCT 464
Db 147 ValTyrIleTyrArgAsnIleLysAspValIleValSerPheTyrHisPheArgGlu 166
Qy 465 CTGGCGGACCATGAGCTACCGA---GGCACCTTTCAAGAAATTCCTCGCGAGAGTTTATGAAT 521
Db 167 IleThrLysLeuProLeuGluAspAlaProPheGluGluAlaPheAspGluPheTyrHis 186
Qy 522 GATAAGCTGGCTACGGCTCTGTGTTTGAGCACGTCGAGGAGTTCTGGGAGCACCGCATG 581
Db 187 GlyLysSerGlnPheGlyProTyrTrpAspHisLeuLeuGlyTyrTrpLysAlaSerLeu 206
Qy 582 GAC-----TCGAACGCTGCTTTTCTCAAGTATGAAGACATGCATCGGAGCCTGGTGACG 635
Db 207 GluArgProGluValIleLeuPheLeuLysTyrGluAspValLysLysAspProThrSer 226
Qy 681 GCCCAGCTGGAAGCCCTGACGAGCAGCTGC-----CACCAGCTGGTG 722
Db 247 GlyValIleGluSerIleIleLysLeuCySerPheGluAsnLeuSerAsnLeuGluVal 266
Qy 723 GACCAGTCTGCAACGCTGAGGCC---CTGCCCGTG-----GGCCGG 761
Db 267 AsnLysSerGlyAsnSerLysGlyPheLeuProIleGluAsnArgLeuTyrPheArgLys 286
Qy 762 GGAAGAGTTGGCTGTGAGGACATCTTACCCTCTCCATGAATGAGAAGTTTGCATTG 821
Db 287 AlalysAspGlyAspTrpLysAsnTyrPheThrAspGluMetThrGluLysIleAspLys 306
Qy 822 GTGTATAACAGAAAGATGGAAGTGTGACCTC 854
Db 307 LeuIleAspGluLysLeuSerAlaThrGlyLeu 317
RESULT 91
US-11-096-568A-12566
; Sequence 12566, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 12566
; LENGTH: 340
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(340)
; OTHER INFORMATION: Ceres Seq. ID no. 14302318
US-11-096-568A-12566
Alignment Scores:
Pred. No.: 7.59e-14 Length: 340
Score: 275.00 Matches: 95
Percent Similarity: 41.7% Conservative: 49
Best Local Similarity: 27.5% Mismatches: 127
Query Match: 6.2% Indels: 74
DB: 6 Gaps: 13
US-10-768-158-1 (1-2419) x US-11-096-568A-12566 (1-340)
Qy 3 GACCGCCACGGC-----GGCGCATGGCGGAGAGCGAGGCC 38
Db 9 AspGlyAspGlyProHisAspGlyAlaProThrTyrAlaGlyMetAlaLysLeu----- 26


```
Qy 39 GAGACCCACACCCCGGGGAGTTCGAGAGCAAGTACTTCGAGTTCCTCCATGGCGTGGG 98
Db 27 ---ValProAlaLeuProLeuGlu-----ThrArg 35
Qy 99 CTGCGCGCCTCTCTGC----- 113
Db 36 CysProProPheAlaLeuArgArgTyrGlyGlyPheTrpLeuProGluAlaLeuArg 55
Qy 114 CGCGGGAAGATGGAGGAGATGCCCAACTTCCCGTGGCGCCACGACGCTGTGGATGTC 173
Db 56 ArgGlyLeuProAlaLeuHisAlaGlyPheAlaProArgArgAlaAspValLeuLeuAla 75
Qy 174 ACCTACCCCAAGTCGCGCACACGCTTGTGAGAGGTGCTACTTGTGTCAGCCAGGC 233
Db 76 SerTyrProLysSerGlyThrTrpLeuLysAlaLeuAlaPheAlaThrValArgArg 95
Qy 234 GCT-----GACCCCGATGAGATCGGC-----TTGATGAACATCGACGACGAG 275
Db 96 AlaValHisProProSerSerIleAspHisProLeuArgArgArgAsnProGluAspCys 115
Qy 276 CTCCCGGTCTCGAGTACCCACAGCCCGGCGCTGGACATC-----ATCAAGGAACAGC 329
Db 116 ValProPheLeuGluPheAspPheAspHisGlyAspMetAlaGluLeuGluAlaLeuPro 135
Qy 330 TCTCCCGCCTCATCAGAGCCACTGCGCCTACCGCTTCTGCGCTCTGACCTCCACAAT 389
Db 136 SerProArgValLeuAlaThrHisLeuProTyrSerLeuLeuProAspArgValVala 155
Qy 390 GGAGAC---TCCAAAGTCACTATATGCTCGCAACCCCAAGGATCTGGTGTCTTAT 446
Db 156 AspAspGlyAlaArgIleValTyrValCysArgAspProLysAspAlaLeuValSerSer 175
Qy 447 TATCAGTTCCACCG-----TCTCTGCGGACCATGAGCTACCGAGGACCTTT 494
Db 176 TrpLeuPheThrArgLysAlaSerProSerValGlyValAspSerAspSerTyrSerLeu 195
Qy 495 CAAGAATCTCCCGAGGTTTATGAATGATGAAGTGGCTACGGCTCTGTTGTGAGCAC 554
Db 196 GlnGluAlaAlaGluLeuPheArgGluGlyArgCysPheTyrGlyProGlnCysArgHis 215
Qy 555 GTGAGGAGTCTGGGACACCGCATG-----GACTCGAAGTCTCTTTTCTC 602
Db 216 ValLeuGluTyrTrpAspAlaSerArgArgProPheGlyGlyGlyGluValLeuPheLeu 235
Qy 603 AAGTATGAAGACATCGATCGGACCTGCTGACGATGTGTGAGCAGCTGCGCAGATCTCTG 662
Db 236 ArgTyrGluGluMetLeuArgAspProAlaGlySerLeuLysThrMetAlaGluPheMet 255
Qy 663 GGGGTGCTGTGACAGCCCGACGCTGAAGCCCTGACGGAGCAGCTGCCACCACTGGTG 722
Db 256 Gly-----CysGlyPheSerGluGluGluValAlaArgGlyValValAspGluIleVal 273
Qy 723 GACCACTGCTGC-----AACGCTCAGGCCCTGCCGTCGGCGCGG----- 761
Db 274 AspLeuCysSerLeuGluLysLeuLysAsnMetGluAlaAsnArgAspGlyArgArgAsn 293
Qy 762 -----GGAAGAGTTGGGCTGTGGAAGGAC 785
Db 294 AlaSerGlyIleArgSerAspSerPhePheArgLysGlyValAlaAlaGlyAspTrpSerSer 313
Qy 786 ATCTTCACCGTCTCCATGAATGAGAGTTTACCTTGGTGATAAACAGAACTGGGAAG 845
Db 314 HisMetSerProGluMetGlyLysMetLeuAspGluAlaValGluGluAlaLeuGlnGly 333
Qy 846 TGTGACCTCACGTTT 860
Db 334 SerGlyPheSerPhe 338
```

RESULT 92

US-11-096-568A-1018

; Sequence 1018, Application US/11096568A

; Publication No. US20060048240A1

GENERAL INFORMATION:

```
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 1018
; LENGTH: 282
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(282)
; OTHER INFORMATION: Ceres Seq. ID no. 15218614
US-11-096-568A-1018
```

Alignment Scores:

Pred. No.:	8,59e-14	Length:	282
Score:	274.00	Matches:	77
Percent Similarity:	44.7%	Conservative:	42
Best Local Similarity:	28.9%	Mismatches:	107
Query Match:	6.2%	Indels:	40
DB:	6	Gaps:	9

US-10-768-158-1 (1-2419) x US-11-096-568A-1018 (1-282)

```
Qy 138 AACTTCCCGTGGCGCCACGACGCTGTGGATCGTCACTACCCCAAGTCCGGCACGAGC 197
Db 6 HisPheGluAlaArgGlyThrAspValLeuLeuValThrMetProLysSerGlyThrThr 25
Qy 198 TTGCTGACGAGGTGCTACTTGTGTCAGCGCGGCGCTACCCCGATGAGATCGGCTTG 257
Db 26 TrpIleLysAlaLeuLeuTyrAlaAlaHisArgThrAspAspThrSerSerSerIle 45
Qy 258 ATG-----AACATCGACGACGAGCTCCGGTCTCGAG----- 290
Db 46 LeuArgGlnLeuAlaSerHisAsnSerHisGlnLeuValProPheLeuGluAlaGlnVal 65
Qy 291 TACCACAGCGCGGCGCTTGACATCATCAAGAACTGACCTCTCCCGCCTCATCAAGAGC 350
Db 66 TyrThrLysAspGlnIleProAspLeuSerSerLeuProAlaProArgLeuPheAlaThr 85
Qy 351 CACTGCGCCTACCGCTTCTGCGCTCTGACCTCCCAATCGAGACTCCAAGGTCATCTAT 410
Db 86 HisIleProAlaGluSerLeuProSerValValAlaSerGlyCysLysValValTyr 105
Qy 411 ATGCTCGCAACCCCAAGGATCTGGTGTCTTATTATCAGTTCCACCGCTCTCTCGCG 470
Db 106 LeuCysArgAspProLysAspCysPheValSerLeuTrpHisPheMetAsnLysPheThr 125
Qy 471 ACCATGAGTACCGAGGACCTTTTCAAGAAATTCGCGGAGGTTTATGAATGATAAGCTG 530
Db 126 ProTrpAspIleAspGluAlaHisGlyArgPheCysGlu----- 138
Qy 531 GGC-----TACGGCTCTGCTTTCAGACAGCTGAGGAGTCTCGGACACCGCATG 581
Db 139 GlyValSerLeuTyrGlyProPheTrpGluHisValLeuSerTyrTrpArgTrpHisVal 158
Qy 582 GAC-----TCGAACGTGCTTTTCTCAAGTATCAAGACATCATCGGACCTGCTGACG 635
Db 159 AspArgProGlnValLeuPheLeuThrTyrGlnGluLeuSerAlaAspProLeuGly 178
Qy 636 ATGCTGAGCAGCTGGCCAGATCTCTGGGGGTGCTGTGACAAAGGCCAGCTCGAAGCC 695
Db 179 GlnLeuArgArgLeuAlaGluPheIleGlyArgProPheThrProGlyGluGlnGluAla 198
Qy 696 -----CTGACGGGACACTGC-----CACCAGCTGCTG 722
Db 199 GlyValAspArgGluIleAlaGluAlaCysAlaMetLysSerMetValAsnGlnGluVal 218
Qy 723 GACCACTGCTGCAACGCTGAG-----GCCCTGCCCGTGGCG----- 758
```


Db 84 -----AspGluHisHisProLeuLeuGluHisAsnProHisAspLeuValGly 99
Qy 297 -----CAGCGGGCTGGACATATCAAGGAACCTGACCTCTCCCGGCTCATC 344
Db 100 PheLeuGluIleTyrProLysLeuGluLeuTyrGluSerLeuProSerProArgLeuLeu 119
Qy 345 AAGAGCCACCTGCGCTACCGCTTCTGCGCTCTGACCTCCCAATGGAGACTCC----- 398
Db 120 SerThrHisLeuProTyrSerMetLeuProHisHisArgGluGluGlnThrGlyCys 139
Qy 399 AAGGTCTATATATGGCTCGCAACCCCAAGGATCTGGTGTCTTATATCACTTCCAC 458
Db 140 ArgLeuValTyrIleTyrArgAspProLysAspAlaMetValSer-----MetTrpHis 157
Qy 459 CGCTCTCGCGACCATGAGCTACCGAGGCACCTTTCAAGATTCTCGCGAGGTTTATG 518
Db 158 GlnAsnLysLysGluLysLysAsnArgLeuThrPheGluGluMetPheAspMetPheCys 177
Qy 519 AATGATAAGCTGGGCTACCGCTCTGCTGTTTCAGCAGCTGCAGGAGTTCTGGGACGACCGC 578
Db 178 GluGlyArgCysValValGlyProGlnTrpCysHisAlaGlyGluTyrTrpAspLysSer 197
Qy 579 ATG-----GACTCGAAGCTGCTTTTCTCAAGTATGAAGACATGCATCGGAGCTGTG 632
Db 198 GlnAlaArgProGluLysValLeuPheLeuMetTyrGluAspLeuLeuGlnAspThrVal 217
Qy 633 ACGATGTGGAGCAGCTGCGCAGATTCCTGGGGGTCTCTGTGCAAGGCCACGCTGAA 692
Db 218 GlyAsnLeuArgThrLeuAlaGluPheMetGly-----CysGlyPheSerArgGlnGlu 235
Qy 693 GCCCTGCGGAGCAGCTGCCACGCTGGTGACCACTGCTGC-----AACGCTGAGGCC 746
Db 236 GluAspAspGlyIleValGlnGlnIleValGluLeuCysSerLeuAsnAsnLeuLysAsn 255
Qy 747 CTGCGCGTGGGC----- 758
Db 256 LeuAsnValAsnLysSerGlyThrThrLeuLeuGlyIleSerLysAspGlyPheArg 275
Qy 759 CGGGGAAGAGTTGGCTGTGAAGACATCTTACCGCTCTCCATGAATGAGAAGTTGAC 818
Db 276 LysGlyGlyThrGlyAspTrpSerAsnHisMetSerProGluMetAlaAlaArgLeuAsp 295
Qy 819 TTGGTGTTAACAAGACATG 839
Db 296 LysIleValLysGluArgLeu 302

RESULT 95

US-11-096-568A-7712
; Sequence 7712, Application US/11096568A
; Publication No. US20060048240A1

GENERAL INFORMATION:

; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 7712
; LENGTH: 329
; TYPE: PPT

ORGANISM: Glycine max

; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(329)
; OTHER INFORMATION: Ceres Seq. ID no. 15174866

US-11-096-568A-7712

Alignment Scores:

Pred. No.: 2,64e-13 Length: 329
Score: 268.50 Matches: 78
Percent Similarity: 44.2% Conservative: 45
Best Local Similarity: 28.1% Mismatches: 106

Query Match: 6.1% Indels: 49
DB: 6 Gaps: 9

US-10-768-158-1 (1-2419) x US-11-096-568A-7712 (1-329)

Qy 141 TTCCCGGTGGCGCCAGCGACGTGTGATCGTCACCTACCCCAAGTCGCGCACCAGCTTG 200
Db 63 PheValProArgAlaAspAspIleIleLeuAlaThrGlnProLysCysGlyThrTrp 82
Qy 201 CTCAGAGGTGGTCTACTTGGTAGCCAG-----GGCGGTGACCCCGATGAG 248
Db 83 LeuLysAlaLeuAlaPheThrIleAlaAsnArgSerArgTyrGlyPheSerAspHisPro 102
Qy 249 ATCGGCTTGATGAACATCGAGCAGCTCCGGTCTCGGAGTACCA-----CAGCCG 302
Db 103 LeuLeuThrArgHisProGlnHisValProPheIleGluIleProGlyAlaGlyPro 122
Qy 303 GGCCTGACATCATCAAGGAACCTCTCCCGCTCATCAAGAGCCACCTGCCCTAC 362
Db 123 AsnHisThrAspIleHisThrLeuProProArgLeuLeuAlaThrHisMetProMet 142
Qy 363 CGTTTCTGCCCTCTGACCTCCACAATGGAGACTCCAAAGTCACTTATATGCTCGCAAC 422
Db 143 SerLeuLeuProProGlyThrArgSerValGlyCysArgValValTyrLeuCysArgAsp 162
Qy 423 CCCAAGGATCTGGTGTCTTATTATCAGTTCCACCGCTCTCTGCGGACCATGAGCTAC 482
Db 163 ProLysAspAlaLeuValSerArgLeuHisPheGluAsnLys-----AlaPhe 178
Qy 483 CGAGGCACC-----TTTCAAGAAATTCGCGGAGGTTTATGAAT 521
Db 179 GlnGlyThrAsnLeuSerMetAspAsnAlaPheSerMetPheCysGluGlyPheSer--- 197
Qy 522 GATAAGCTGGCTACGCTCTCTGTTGACACGTGCAGGAGTTCTCGGAGCACCGCATG 581
Db 198 -----ProTyrGlyProPheTrpAspHisCysLeuGlyTyrTrpArgGluSerIle 214
Qy 582 -----CACTCGACGTGCTTTTCTCAAGTATGAACATGCATCGGACCTGTCACG 635
Db 215 AlaArgProAspAsnValLeuPheLeuLysTyrGluIleLysSerAspProValPro 234
Qy 636 ATGTGTGAGCAGCTGGCGAGATTCTCGGGGTGCTCTGCACAGGCCAGCTGGAAGCC 695
Db 235 ValValArgLysLeuAlaLysPheLeuGlyValProLeuThrGlu-----GluGlu 251
Qy 696 CTGACCGGACACTCCACACGCTGGTGGACACGTGCTGCAACGCTGAGGCCCTG----- 749
Db 252 GluSerSerGlyValAlaGlnGluValValLysLeuCysSerPheGluThrLeuThrSer 271
Qy 750 -----CCCGTGGGC----- 758
Db 272 LeuGlnValAsnGlnValGlyValArgHisGlyAspLysIlePheMetSerAsnSer 291
Qy 759 -----CGGGGAAGAGTTGGCTGTGGGAAGGACATCTTACCGCTCTCCATGAAT 806
Db 292 ValPheTyrArgLysGlyGluValGlyAspTrpAlaAsnHisMetSerHisValMetGly 311
Qy 807 GAGAAGTTTGACTTGGTGTATAAACAGAAAGTGGAAAGTGTGACCTCACCTTT 860
Db 312 GluGluLeuAspArgIleValGlnGlnLysLeuGlySerGlyLeuValPhe 329

RESULT 96

US-11-096-568A-7711
; Sequence 7711, Application US/11096568A
; Publication No. US20060048240A1

GENERAL INFORMATION:

; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471

```
; SEQ ID NO 7711
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(330)
; OTHER INFORMATION: Ceres Seq. ID no. 15174865
US-11-096-568A-7711

Alignment Scores:
Pred. No.: 2,65e-13 Length: 330
Score: 268.50 Matches: 78
Percent Similarity: 44.2% Conservative: 45
Best Local Similarity: 28.1% Mismatches: 106
Query Match: 6.1% Indels: 49
DB: Gaps: 9

US-10-768-158-1 (1-2419) x US-11-096-568A-7711 (1-330)
Qy 141 TTCGGGTGGCCGAGCGAGCTGTGGATCTCCTACCTACCCCAAGTCGGGACCGAGCTTG 200
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
64 PheValProArgAlaAspAspIleLeuAlaThrGlnProLysCysGlyThrTrp 83
Qy 201 CTGCAGGAGGTGCTACTCTGTGAGCCAG-----GGCGTGCACCCCGATGAG 248
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
84 LeuLysAlaLeuAlaPheThrIleAlaAsnArgSerArgTyrGlyPheSerAspHisPro 103
Qy 249 ATCGGCTTGATGAACATCGAGCAGCTCCCGCTCTGGAGTACCCCA-----CAGCG 302
Db :::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
104 LeuLeuThrArgHisProGlnHisValProPheIleGluIleProGlyAlaGlyPro 123
Qy 303 GGCTTGACATCATCAAGAACTGACCTCTCCCGCTCATCAAGAGCCACTGCCCTAC 362
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
124 AsnHisThrAspIleHisThrLeuProProArgLeuLeuAlaThrHisMetProMet 143
Qy 363 CGCTTTCTGCCTCTGACCTCCACATGGAGACTCCCAAGGTCTATATATGGCTCGCAC 422
Db :::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
144 SerLeuLeuProProGlyThrArgSerValGlyCysArgValValTyrLeuCysArgAsp 163
Qy 423 CCCAAGGATCGTGGTGTCTATTATCAGTTTCACCGCTCTCTCGGACCATCAGCTAC 482
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
164 ProLysAspAlaLeuValSerArgLeuHisPheGluAsnLys-----AlaPhe 179
Qy 483 CGAGGCACC-----TTTCAAGAAATCTCGCGAGGTTTATGAAT 521
Db ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
180 GlnGlyThrAsnLeuSerMetAspAsnAlaPheSerMetPheCysGluGlyPheSer--- 198
Qy 522 GATAAGCTGGCTACGGCTCTCGTTGAGCAGCTGCGAGGAGTCTCGGACACCGCATG 581
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
199 -----ProTyrGlyProPheTrpAspHisCysLeuGlyTyrTrpArgGluSerIle 215
Qy 582 -----GACTCGAAGCTGCTTTTCTCAAGTATGAAGATCATCGGACCTGGTGACG 635
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
216 AlaArgProAspAsnValLeuPheLeuLysTyrGluGluIleLysSerAspProValPro 235
Qy 636 ATGTGGAGCAGCTGGCCAGATTCCTGGGGGTGCTCTGTGACAAGGCCCGCAGCTGGAAGCC 695
Db ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
236 ValValArgLysLeuAlaLysPheLeuGlyValProLeuThrGlu-----GluGlu 252
Qy 696 CTGACGGAGCACTGCCACCACTGGTGACCAAGTGTGCAACCGTGGAGCCCTG----- 749
Db ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
253 GluSerSerGlyValAlaGlnGluValValLysLeuCysSerPheGluThrLeuThrSer 272
Qy 750 -----CCGTGGGC----- 758
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
273 LeuGlnValAsnGlnValGlyGlyValArgHisGlyAspLysIlePheMetSerAsnSer 292
Qy 759 -----CGGGGAAGAGTGGGTGGAGGACATCTTCACCGTCTCCATGAAT 806
Db ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
293 ValPheTyrArgLysGlyGluValGlyAspTrpAlaAsnHisMetSerHisValMetGly 312
Qy 807 GAGAAGTTTGACTTGGTGTATAAACAAGATGGAAAGTGTGACCTCACGTTT 860
```

```
Db 313 GluGluLeuAspArgIleValGlnGlnLysLeuCluclySerGlyLeuValPhe 330
RESULT 97
US-11-096-568A-7710
; Sequence 7710, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 7710
; LENGTH: 340
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(340)
; OTHER INFORMATION: Ceres Seq. ID no. 15174864
US-11-096-568A-7710

Alignment Scores:
Pred. No.: 2,68e-13 Length: 340
Score: 268.50 Matches: 78
Percent Similarity: 44.2% Conservative: 45
Best Local Similarity: 28.1% Mismatches: 106
Query Match: 6.1% Indels: 49
DB: Gaps: 9

US-10-768-158-1 (1-2419) x US-11-096-568A-7710 (1-340)
Qy 141 TTCGGGTGGCCGAGCGAGCTGTGGATCTCCTACCTACCCCAAGTCGGGACCGAGCTTG 200
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
74 PheValProArgAlaAspAspIleLeuAlaThrGlnProLysCysGlyThrTrp 93
Qy 201 CTGCAGGAGGTGCTACTCTGTGAGCCAG-----GGCGTGCACCCCGATGAG 248
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
94 LeuLysAlaLeuAlaPheThrIleAlaAsnArgSerArgTyrGlyPheSerAspHisPro 113
Qy 249 ATCGGCTTGATGAACATCGAGCAGCTCCCGTCTCGGAGTACCCCA-----CAGCG 302
Db ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
114 LeuLeuThrArgHisProGlnHisValProPheIleGluIleProGlyAlaGlyPro 133
Qy 303 GGCTTGACATCATCAAGAACTGACCTCTCCCGCTCATCAAGAGCCACTGCCCTAC 362
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
134 AsnHisThrAspIleHisThrLeuProProArgLeuLeuAlaThrHisMetProMet 153
Qy 363 CGCTTTCTGCCTCTGACCTCCCAATGGAGACTCCCAAGGTCACTATATGGCTCGCAAC 422
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
154 SerLeuLeuProProGlyThrArgSerValGlyCysArgValValTyrLeuCysArgAsp 173
Qy 423 CCCAAGGATCTGGTGTCTTATTATCAGTTTCACCGCTCTCTCGGACCATCAGCTAC 482
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
174 ProLysAspAlaLeuValSerArgLeuHisPheGluAsnLys-----AlaPhe 189
Qy 483 CGAGGCACC-----TTTCAAGAAATCTCGCGAGGTTTATGAAT 521
Db ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
190 GlnGlyThrAsnLeuSerMetAspAsnAlaPheSerMetPheCysGluGlyPheSer--- 208
Qy 522 GATAAGCTGGCTACGGCTCTCTGGTTGAGCAGCTGCGAGGATCTCGGACACCGCATG 581
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
209 -----ProTyrGlyProPheTrpAspHisCysLeuGlyTyrTrpArgGluSerIle 225
Qy 582 -----GACTCGAAGCTGCTTTTCTCAAGTATGAAGATCATCGGACCTGGTGACG 635
Db ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
226 AlaArgProAspAsnValLeuPheLeuLysTyrGluGluIleLysSerAspProValPro 245
Qy 636 ATGTGGAGCAGCTGGCCAGATTCCTGGGGGTGCTCTGTGACAAGGCCCGCAGTGAAGCC 695
Db ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Db 246 ValValArgLysLeuAlaLysPheLeuGlyValProLeuThrGlu-----GluGlu 262
Qy 696 CTGACGAGACTGCCACAGCTGGTGACAGTGCTGCAACGCTGAGGCCCTG----- 749
Db 263 GluSerSerGlyValAlaGlnGluValValLysLeuCysSerPheGluThrLeuThrSer 282
Qy 750 -----CCCGTGGGC----- 758
Db 283 LeuGlnValAenGlnValGlyValArgHisGlyAspLysIlePheMetSerAsnSer 302
Qy 759 -----CGGGAGAGTTGGCTGTGGAGGACATCTTCACCGTCTCOATGAAT 806
Db 303 ValPheTyArgLysGlyGluValGlyAspTrpAlaAenHisMetSerHisValMetGly 322
Qy 807 GAGAAGTTGACTGGTGTTATAACACAGAGATGGAAAGTGTGACCTCACGTTT 860
Db 323 GluGluLeuAspArgIleValGlnGlnLysLeuGluGlySerGlyLeuValPhe 340

RESULT 99
US-10-767-701-42402
; Sequence 42402, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 42402
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C44289_1.pap
US-10-767-701-42402

Alignment Scores:
Pred. No.: 3,26e-13 Length: 344
Score: 267.50 Matches: 78
Percent Similarity: 45.5% Conservative: 47
Best Local Similarity: 28.4% Mismatches: 109
Query Match: 6.1% Indels: 41
DB: 4 Gaps: 10

US-10-768-158-1 (1-2419) x US-10-767-701-42402 (1-344)

Qy 141 TTCGGGTGGCGCCAGCAGCTGTGGATCGTCACTACCCCAAGTCCGGACACGAGTTG 200
Db 75 PheThrProArgProAspAspValIleLeuAlaSerTyProLysCysGlyThrThrTrp 94
Qy 201 CTGACGAGGTGGTCTACTTGTGTGAGC-----CAGGCGCTGACCCCGATGAGATCGGC 254
Db 95 LeuLysAlaLeuAlaPheAlaThrAlaAlaArgThrAlaTyProProSerAspAlaGly 114
Qy 255 -----TTGATCAACATCAGCAGCAGCTCCCGCTCGGTGAGTACCCACAGCCG 302
Db 115 HisProLeuArgArgMetAsnProHisAspCysIleProPheIleAspGluIlePheAla 134
Qy 303 GGCCTGGACATCATCAAG-----GAACGTACCTTCCCGCTCATCAAGACCACTG 356
Db 135 GlyGlyGluAspAlaLysLeuAspAlaLeuProSerProArgLeuMetAsnThrHisLeu 154
Qy 357 CCCTACCGCTTCTGCGCTCTGACCTC-----CACAAATGGAGCTCCCAAGTTCATCTAT 410
Db 155 ProTyAlaLeuProProIleSerValValThrAlaThrGlyGlyCysArgValAlaTy 174
Qy 411 ATGGCTCGCAACCCCAAGATCTGGTGTCTTATTATCATGTTCC---CACCGTCTCTG 467
Db 175 ValCysArgAspProLysAspMetValValSerLeuTrpHisPheLeuArgArgThrLys 194
```

```
Qy 468 CGACCATGAGCTACCGAGGCACCTTTCAAGAAATTCCTCGCGAGGTTTATGAATGATAAG 527
Db 195 ProAspLeuSerPheAlaAspThrPheGluSerValCys-----AspGlyThr 210
Qy 528 CTGGGCTACCGCTCTCGTTTGTAGCAGCTGCAGAGTTCTGGGAGCACCGCATG----- 581
Db 211 ValAlaAlaGlyProValTrpAspHisValLeuSerTyTrpArgAlaSerValAlaThr 230
Qy 582 GACTCGAACCTGCTTTTCTCAAGTATGAAGACATGATCGGACCTGGTGACATGGTG 641
Db 231 ProAspArgAlaLeuPheLeuLysTyArgAspLeuLeuArgAspProAlaGlyAsnVal 250
Qy 642 GAGCAGCTGCCGATTTCTCTGGGGGTCTCTGTGACAAGGCCACGAGTGGAGCCCTGACG 701
Db 251 ArgArgLeuAlaGluPheMetGly-----CysProPheSerAlaAlaGluGluAla 268
Qy 702 GAGCAGCTGCCACGAGCTGTGGACAGTGTGCAACGCTGAG-----GTG 755
Db 269 GlyAsnValAlaAlaValValGluLeuCysSerPheAspGluMetLysGlyLeuGluVal 288
Qy 744 -----GCCCTGCCC-----GTG 755
Db 289 AsnArgProGlyGlyThrAlaGlyLysTyArgAlaMetProArgAspAlaPhePhe 308
Qy 756 GGCCGGGAGAGTTGGCTGTGGAAGGACATCTTCAACCGTCTCCATGAATGAGAAGTTT 815
Db 309 ArgLysGlyValAlaGlyAspTrpAlaAsnHisMetThrProGluMetAlaAlaArgLeu 328
Qy 816 GACTTGTGTATAAACAGAAGATGGGAAAGTGTGACCTCACGTTT 860
Db 329 AspGlnIlePheArgAspLysLeuGlnGlyThrGlyLeuAlaPhe 343
```

```
RESULT 99
US-10-259-194A-330
; Sequence 330, Application US/10259194A
; Publication No. US20040010815A1
; GENERAL INFORMATION:
; APPLICANT: Lange, Markus B.
; APPLICANT: Ghasssemlan, Majid
; APPLICANT: Briggs, Steven P.
; APPLICANT: Cooper, Bret
; APPLICANT: Glazebrook, Jane
; APPLICANT: Goff, Stephen A.
; APPLICANT: Katagiri, Fumiyaki
; APPLICANT: Kreps, Joel
; APPLICANT: Moughamer, Todd
; APPLICANT: Provart, Nicholas
; APPLICANT: Ricke, Darrell
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: IDENTIFICATION AND CHARACTERIZATION OF PLANT GENES
; FILE REFERENCE: 70029-NP
; CURRENT APPLICATION NUMBER: US/10/259,194A
; CURRENT FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/370,743
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 662
; SOFTWARE: PatentList.pl version 3.0.4 (C) 2001 Syngenta
; SEQ ID NO 330
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-259-194A-330

Alignment Scores:
Pred. No.: 3,59e-13 Length: 343
Score: 267.00 Matches: 91
Percent Similarity: 44.2% Conservative: 58
Best Local Similarity: 27.0% Mismatches: 122
```

Query Match:	6.1%	Indels:	66
DB:	4	Gaps:	16
US-10-768-158-1 (1-2419) x US-10-259-194A-330 (1-343)			
QY	9	GACGGCGCGCATCGCGGAGACGCGCGGAGACCCCGACACCCCGCGGGAGTTC---	65
DB	16	AspAspGlyThrLeuAlaAlaGluAlaArgSerProLeuProLysGluGluPheGly	35
QY	66	-----GAGACCAAGTACTTCGAG-----TTCCATGCG	92
DB	36	AspLeuValAlaAlaLeuProArgLysGluGlnTyrLeuAspGlyArgLeuTyrGluGly	55
QY	93	GTGGCGTCCGCCCC---TTCGCGCGGGAGATGGAGAGATCCCAACTTCGCGGTG	149
DB	56	PheTrpLeuProGluHisTyrAlaProGlyIleAlaPheArgArgPheTrpPro	75
QY	150	CGGCCACGCGTGTGATCGTCACTACCCCAAGTCGGCACCAAGCTTGTGCAGGAG	209
DB	76	ArgAlaAspValValLeuAlaSerTyrProLysCysGlyThrTrpLeuLysAla	95
QY	210	GTGCTCTAC---TTGGTGAGCGAGCGCGTGCACCC---GATGAGATCGGTTGATGAAC	263
DB	96	LeuAlaPheAlaAlaMetThrArgAlaAlaTyrProAlaAspGluHisProLeuLeuArg	115
QY	264	ATCGAC-----GACGAGCTCCCGTCTCGGAGTACCCACACCGGGCCCTGGACATCATC	317
DB	116	LeuAenProHisAspValIleProPheValGlu-----AspValPhe	129
QY	318	AAGGAA-----CTGACCTCTCCCGCTCATCAAGAGCCAC	353
DB	130	ThrAspGlyHisGluAlaLysLeuAspMetLeuProSerProArgLeuIleAsnThrHis	149
QY	354	CTGCCCTACCGCTTCTGCCCTCTGACCTCCACAATGGAGACTCC-----AAG	401
DB	150	ThrProTyrGlnLeuLeuProGluSerValValAlaGlyAspGlyGlyGlyCysLys	169
QY	402	GTCATCATATGGCTCGCAACCCCAAGGATCTGGTGTGTCTTATATCAGTTCACCGC	461
DB	170	ValValTyrIleCysArgAspProLysAspMetValValSerLeuTyrHisPheMetArg	189
QY	462	TCTCTGCGG---ACCATGAGTCTACCGAGCAGCTTTCAGAAATTTCTCGCGGAGTTTATG	518
DB	190	ArgLeuGlnProAspLeuSerLeuAlaGlyValValGluSerValAla-----	205
QY	519	AATGATAGCTGGGCTACGGCTCTGGTTTGACGACGTGTCAGAGGTTCTGGGAGCACCGC	578
DB	206	AspGlyThrValProPheGlyProMetTrpAspHisIleLeuGlyTyrTrpArgAlaSer	225
QY	579	ATGGAC-----TCGAAACGTCTTTTCTCAAGTATGAACATGCATCGGACCTGGTG	632
DB	226	ValSerArgProAspArgValLeuLeuLeuArgTyrGluAspLeuLeuArgAspGlyAla	245
QY	633	ACG-----ATGTGTGGACGATGCGCCAGATTCCTTGGGGGTCTCTGTGACAAAGGCCAC	686
DB	246	AlaGlyGluHisValArgAlaMetAlaArgPheMetGlyArgProPheSerAlaAla---	264
QY	687	CTGAAGCCCTGACGGAGACCTGCCACGACGTGTGGACCAAGTGTCTGTC-----AACGCT	740
DB	265	---GluGluAlaAlaGlyAlaValAlaSerValValGluLeuCysSerPheGluArgMet	283
QY	741	GAGGCTCTCCCGTGGGCGG-----	761
DB	284	LysAlaLeuGluValAsnArgArgGlyThrAlaGlySerTyrLysSerMetProArgAsp	303
QY	762	-----GGAAGAGTGTGGGTGTGGGAAGGACATCTTCACCGTCTCCATGAAT	806
DB	304	AlaPhePheArgLysGlyValAlaGlyAspTrpAlaAsnHisMetSerProGluThrAla	323
QY	807	GAGAAGTTTGATCTGGTGTATAACAGAAGATGGGAAAGTGTGACCTACG	857
DB	324	AlaArgLeuAspGlyIlePheArgGluLysPheArgGlyThrGlyLeuThr	340

```

RESULT 100
US-10-437-963-159374
; Sequence 159374, Application US/10437963
; Publication No. US2004012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 159374
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_58757C.1.pep
US-10-437-963-159374

Alignment Scores:
Pred. No.:      3,59e-13      Length:      343
Score:          267.00      Matches:      91
Percent Similarity: 44.2%      Conservative: 58
Best Local Similarity: 27.0%      Mismatches: 122
Query Match:      6.1%      Indels:      66
DB:              4      Gaps:      16

US-10-768-158-1 (1-2419) x US-10-437-963-159374 (1-343)

Qy 9 GACGGCGGGCATGGCGGAGACGCGGCGGAGACCCCGACCCCGGGGGAGTTC--- 65
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16 AspAspGlyThrLeuAlaAlaAlaGluAlaA-gSerProLeuProLysGluGluPheGly 35
   66 -----GAGACCAAGTACTTCGAG-----TTCATGGC 92
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 36 AspLeuValAlaAlaLeuProHArgLysGluGlnTyLeuAspGlyArgLeuTyGluGly 55
   93 GTGGCGGTGCGCGCC---TTCTGCGCGGGAGATGGAGGAGATCGCAACTTCCCGGTG 149
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 56 PheTrpLeuProGluHisTyAlaProGlyIleAlaPheArgArgPheThrPro 75
   150 CGGCCCCAGCGAGTGTGGATCTCACTTACCCGAGTCGGCAGCACCGAGTTCGCGAGGAG 209
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 76 ArgAlaAspAspValValLeuAlaSerTyProLysCysGlyThrThrTrpLeuLysAla 95
   210 GTGGTCTAC---TTGGTGAGCCAGGCGCTGACGCC---GATGAGATCGGCTTGATGAAC 263
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 96 LeuAlaPheAlaAlaMetThrArgAlaAlaTyProAlaAspGluHisProLeuLeuArg 115
   264 ATCGAC-----GAGCAGCTCCGGTCTCTGGAGTACCCACAGCCGGCGCTGGACATCATC 317
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 116 LeuAsnProHisAspValIleProPheValGlu-----AspValPhe 129
   318 AAGGAA-----CTGACCTCTCCCGCCTCATCAAGAGCCAC 353
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 130 ThrAspGlyHisGluAlaLysLeuAspMetLeuProSerProArgLeuIleAsnThrHis 149
   354 CTGGCCTACCGCTTTCTGCCCTCTGACCTTCCCAATGGAGAGTCC-----AAG 401
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 150 ThrProTyGlnLeuLeuProGluSerValValAlaGlyAspGlyGlyGlyCysLys 169
   402 GTCATCTATATGGTCGCAACCCCAAGGATCTGGTGGTGTCTTATTATCAGTTCCACCGC 461
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 170 ValValTyIleCysArgAspProLysAspMetValValSerLeuTyHisPheMetArg 189

```


c 85 165.5 3.8 306 2 US-09-252-991A-25699 Sequence 25699, A
 c 86 165.5 3.8 511 2 US-09-252-991A-26078 Sequence 26078, A
 c 87 164.5 3.7 218 2 US-09-252-991A-24321 Sequence 24321, A
 c 88 164.5 3.7 315 2 US-09-252-991A-20553 Sequence 20553, A
 c 89 164 3.7 288 2 US-09-252-991A-32807 Sequence 32807, A
 c 90 163.5 3.7 341 2 US-09-252-991A-32424 Sequence 32424, A
 c 91 163 3.7 233 2 US-09-252-991A-18455 Sequence 18455, A
 c 92 163 3.7 246 2 US-09-252-991A-16586 Sequence 16586, A
 c 93 163 3.7 248 2 US-09-252-991A-22391 Sequence 22391, A
 c 94 162.5 3.7 663 2 US-09-252-991A-30843 Sequence 30843, A
 c 95 161.5 3.7 247 2 US-09-252-991A-26899 Sequence 26899, A
 c 96 161.5 3.7 272 2 US-09-252-991A-31371 Sequence 31371, A
 c 97 160.5 3.6 333 2 US-09-252-991A-28132 Sequence 28132, A
 c 98 160.5 3.6 663 2 US-09-252-991A-30843 Sequence 30843, A
 c 99 160 3.6 208 2 US-09-252-991A-32166 Sequence 32166, A
 c 100 159.5 3.6 195 2 US-09-252-991A-30009 Sequence 30009, A

ALIGNMENTS

RESULT 1

US-09-786-240-7
 ; Sequence 7, Application US/09786240
 ; Patent No. 6558935
 ; GENERAL INFORMATION:
 ; APPLICANT: INCYTE PHARMACEUTICALS, INC.
 ; APPLICANT: TANG, Y. Tom
 ; APPLICANT: CORLEY, Neil C.
 ; APPLICANT: GUEGLER, Karl J.
 ; APPLICANT: BAUGHN, Mariah R.
 ; APPLICANT: LAL, Preeti
 ; APPLICANT: YUE, Henry
 ; APPLICANT: HILLMAN, Jennifer L.
 ; APPLICANT: AZIMZAI, Yalda
 ; TITLE OF INVENTION: HUMAN TRANSFERASE PROTEINS
 ; FILE REFERENCE: PF-0592 PCT
 ; CURRENT APPLICATION NUMBER: US/09/786,240
 ; CURRENT FILING DATE: 2002-03-12
 ; PRIOR FILING DATE: 1998-09-10; 1998-11-04; 1999-05-11
 ; PRIOR APPLICATION NUMBER: 09/150,657; unassigned; 60/133,642
 ; PRIOR FILING DATE: 1998-09-10; 1998-11-04; 1999-05-11
 ; NUMBER OF SEQ ID NOS: 33
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 7
 ; LENGTH: 284
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: Incyte ID No. 6558935 1784742CD1
 US-09-786-240-7

Alignment Scores:
 Pred. No.: 5.32e-142 Length: 284
 Score: 1530.00 Matches: 284
 Percent Similarity: 100.0% Conservat: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 34.8% Indels: 0
 DB: 2 Gaps: 0

US-10-768-158-1 (1-2419) x US-09-786-240-7 (1-284)

Qy 21 ATGGCGGAGCGAGCGGCGAGACCCCGAGCCCGGGGGAGTTTCGAGAGCAAGTACTTC 80
 Db 1 MetAlaGluSerGluAlaGluThrProSerThrProGlyGluPheGluSerLysThrPhe 20
 Qy 81 GAGTTCATGCGTGGCGGCTCCGCGCTTCGCGCGGGAAGATGGAGAGATCGCCCAAC 140
 Db 21 GluPheHisGlyValArgLeuProPheCysArgGlyLysMetGluGluAlaAsn 40
 Qy 141 TTCGCGTGGCGGCGGAGCGAGTGTGGATCGTACCTACCCCAAGTCGGGACAGCTTG 200
 Db 41 PheProValArgProSerAspValTrpIleValThrTyProLysSerGlyThrSerLeu 60

Qy 201 CTGCGAGGAGTGGTCTACTTTGGTGAGCGAGCGGCTCACCCCGATGAGATCGGCTTGATG 260
 Db 61 LeuGlnGluValValTyLeuValSerGlnGlyAlaAspProaspGluIleGlyLeuMet 80
 Qy 261 AACATCGACGAGCAGCTCCCGGTCTCTGAGTACCCACAGCGGGCCTGGACATCATCAAG 320
 Db 81 AsnIleAspGluGlnLeuProValLeuGluTyProGlnProGlyLeuAspIleLeuLys 100
 Qy 321 GAATGACCTCTCCCGCCTCATCAAGAGCCACTGCGCCTACCGCTTCTGCGCTCTGAC 380
 Db 101 GluLeuThrSerProArgLeuIleLysSerHisLeuProTyArgPheLeuProSerAsp 120
 Qy 381 CTCCACAATCGAGACTCCAAAGTCACTATATGCTCGCAACCCCAAGGATCTGGTGGTG 440
 Db 121 LeuHisAsnGlyAspSerLysValIleTyMetAlaArgAsnProLysAspLeuValVal 140
 Qy 441 TCTTATTATCAGTTCCACCGCTCTCTCGGACCATGAGCTACCGAGCACCTTTCAAGAA 500
 Db 141 SerTyTyGlnPheHisArgSerLeuArgThrMetSerTyArgGlyThrPheGlnGlu 160
 Qy 501 TTCTCCCGGAGGTTTATGAATGATAAGCTGGGCTACGGCTCTGGTTTGACACGTCGAG 560
 Db 161 PheCysArgArgPheMetAsnAspLysLeuGlyTyrgLysSerTrpPheGluHisValGln 180
 Qy 561 GAGTTCTGGGAGCACCGCATGGACTCGAACGTGCTTTTCTCAAGTATGAAGACATGCAT 620
 Db 181 GluPheTrpGluHisArgMetAspSerAsnValLeuPheLeuLysTyArgGluAspMetHis 200
 Qy 621 CGGACCTCGTGACGATGGTGGAGCAGCTGCGCAGATTCCTGGGGGTGCTCTGTGACAAG 680
 Db 201 ArgAspLeuValThrMetValGluGlnLeuAlaArgPheLeuGlyValSerCysAspLys 220
 Qy 681 GCCAGCTGGAAGCCCTGACGGAGCACTGCCACGAGCTGTGGACAGTGTGCAACGCT 740
 Db 221 AlaGlnLeuGluAlaLeuThrGluHisCysHisGlnLeuValAspGlnCysCysAsnAla 240
 Qy 741 GAGCCCTCGCGTGGCGGCGGGAAGAGTTGGGCTGTGAAGGACATCTTCACCGTCTCC 800
 Db 241 GluAlaLeuProValGlyArgGlyArgValGlyLeuTrpLysAspIlePheThrValSer 260
 Qy 801 ATGAATCAGAAGTTGACTTGGTGTATATAACAGAAAGATGGAAAGTGTGACCTCACCCTT 860
 Db 261 MetAsnGluLysPheAspLeuValTyLysGlnLysMetGlyLysCysAspLeuThrPhe 280
 Qy 861 GACTTTTATTATA 872
 Db 281 AspPheTyLeu 284

RESULT 2
 US-09-854-122-21
 ; Sequence 21, Application US/09854122
 ; Patent No. 6841718
 ; GENERAL INFORMATION:
 ; APPLICANT: ALBERTE, RANDALL S.
 ; APPLICANT: SMITH, ROBERT
 ; TITLE OF INVENTION: TRANSGENIC PLANTS INCORPORATING TRAIT OF ZOSTERA MARINA
 ; FILE REFERENCE: PHA-007.01
 ; CURRENT APPLICATION NUMBER: US/09/854,122
 ; CURRENT FILING DATE: 2001-09-10
 ; PRIOR APPLICATION NUMBER: 60/202,529
 ; PRIOR FILING DATE: 2000-05-10
 ; NUMBER OF SEQ ID NOS: 51
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 21
 ; LENGTH: 295
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-854-122-21

Alignment Scores:
 Pred. No.: 1.62e-37 Length: 295
 Score: 472.00 Matches: 100
 Percent Similarity: 54.5% Conservative: 52

[illegible]

```
Qy 804 AATGAGAAGTTTGACTTGGTGTATATAACAGAGATGGAAAGTGTGACCTCACGTTT 860
|||||:|||||
Db 273 AsnGluArgPheAspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe 291
|||||:|||||
RESULT 4
US-09-949-016-6018
; Sequence 6018, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6018
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6018
Alignment Scores:
Pred. No.: 3,04e-35 Length: 296
Score: 449.00 Matches: 94
Percent Similarity: 54.6% Conservative: 61
Best Local Similarity: 33.1% Mismatches: 105
Query Match: 10.2% Indels: 24
DB: 2 Gaps: 4
US-10-768-158-1 (1-2419) x US-09-949-016-6018 (1-296)
Qy 72 AAGTACTTCGAGTTCATCGGCTGGGCTCGCGCTCTCTCCGCGGGAAGATGGAGGAG 131
|||||:|||||
Db 12 LysLeuLysGluValGluGlyThrLeuLeuGlnProAlaThrValAspAsnTrpSerGln 31
|||||:|||||
Qy 132 ATCGCCAACTTCCCGTCCGCGCCAGCGAGCTGTGGATCGTCACTACCCCAAGTCGCGC 191
|||||:|||||
Db 32 IleGlnSerPheGluAlaLysProAspPheLeuLeuLeuLeuLeuLeuLeuLeuLeu 51
|||||:|||||
Qy 192 ACCAGCTTGTGTCAGGAGTGTCTACTTGTGTAGCCAGCGCTGACCCCGATGAGATC 251
|||||:|||||
Db 52 ThrThrTrpIleGlnGluIleValAspMetIleGluGlnAsnGlyAspValGluLysCys 71
|||||:|||||
Qy 252 GGCTTGATGAACATCGACGAGCTCCCGCTCTGGAGTAC-----CCACAGCGC 302
|||||:|||||
Db 72 GlnArgAlaIleGlnHieArgHisProPheIleGluTrpAlaArgProProGlnPro 91
|||||:|||||
Qy 303 ---GGCTTGGACATCATCAAGAACTGACCTCTCCGCGCTCATCAAGAGCCACTGCC 359
|||||:|||||
Db 92 SerGlyValGluLysAlaLysAlaMetProSerProArgIleLeuLysThrHisLeuSer 111
|||||:|||||
Qy 360 TACCGCTTCTGCGCTCTGACCTCCACAATGGAGACTCCCAAGGTCACTATATGCTCGC 419
|||||:|||||
Db 112 ThrGlnLeuLeuProProSerPheTrpGluAsnAenCysLysPheLeuTyrValAlaArg 131
|||||:|||||
Qy 420 AACCCCAAGATCTGGTGTCTTATATATCAGTTTCCACCGCTCTCTCGCGACCATGAGC 479
|||||:|||||
Db 132 AsnAlaLysAspCysMetValSerTyrTyrHisPheGlnArgMetAsnHisMetLeuPro 151
|||||:|||||
Qy 480 TACCGAGCACCTTCAAGAACTTCTCGCGAGGTTTATGATGATAGCTGGGTACGGC 539
|||||:|||||
Db 152 AspProGlyThrTrpGluLysPheGluThrPheIleAsnGlyLysValValTrpGly 171
|||||:|||||
Qy 540 TCCTGGTTGAGCAGCTGACGAGTTCCTGGGAGCACCGCATCGGACTCGAAGCTGCTTTT 599
|||||:|||||
```

```
Db 172 SerTrpPheAspHisValLysGlyTrpTrpGluMetLysAspArgHisGlnIleLeuPhe 191
Qy 600 CTCGAAGTATCAAGACATGCGGACCTGGTACCGATGCTGGAGCAGCTGGCCAGATTC 659
|||||:|||||
Db 192 LeuphetyrGluAspIleLysArgAspProLysHisGluIleArgLysValMetGlnPhe 211
|||||:|||||
Qy 660 CTGGGGTGTCTCTGTGACAAAGCCCTGAGCCCTGAGCCCTGACGGAGCAC----- 707
|||||:|||||
Db 212 MetGlyLysLysValAspGluThrValLeuAspLysIleValGlnGluThrSerPheGlu 231
|||||:|||||
Qy 708 -----TGCACACAGCTGTGTGGACGAG 728
Db 232 LysMetLysGluAsnProMetThrAsnArgSerThrValSerLysSerIleLeuAspGln 251
Qy 729 TGCTGCAACGCTGAGGCCCTGCGCGTGGCGGGGAGAGCTTGGCTGTGGAAGGACATC 788
|||||:|||||
Db 252 SerIleSerSerPhe-----MetArgLysGlyThrValGlyAspTrpLysAsnHis 268
|||||:|||||
Qy 789 TTCACCGTCTCCATGAATGAGAAAGTTTGACTTGTGTATATAACAGAGATGGAAAGTGT 848
|||||:|||||
Db 269 PheThrValAlaGlnAsnGluArgPheAspGluIleTyrArgArgLysMetGluGlyThr 288
Qy 849 GACCTCACGTTT 860
Db 289 SerIleAsnPhe 292
RESULT 5
US-09-792-695A-2
; Sequence 2, Application US/09792695A
; Patent No. 7026163
; GENERAL INFORMATION:
; APPLICANT: Freimuth, Robert R.
; APPLICANT: Weinshilboum, Richard M
; APPLICANT: Wieben, Eric D.
; TITLE OF INVENTION: SULFOTRANSFERASE SEQUENCE VARIANTS
; FILE REFERENCE: 07039-273001
; CURRENT APPLICATION NUMBER: US/09/792,695A
; CURRENT FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-792-695A-2
Alignment Scores:
Pred. No.: 3,04e-35 Length: 296
Score: 449.00 Matches: 94
Percent Similarity: 54.6% Conservative: 61
Best Local Similarity: 33.1% Mismatches: 105
Query Match: 10.2% Indels: 24
DB: 3 Gaps: 4
US-10-768-158-1 (1-2419) x US-09-792-695A-2 (1-296)
Qy 72 AAGTACTTCGAGTTCATCGGCTGGGCTCGCGCTCTCTCCGCGGGAAGATGGAGGAG 131
|||||:|||||
Db 12 LysLeuLysGluValGluGlyThrLeuLeuGlnProAlaThrValAspAsnTrpSerGln 31
|||||:|||||
Qy 132 ATCGCCAACTTCCCGTCCGCGCCAGCGAGCTGTGGATCGTCACTACCCCAAGTCGCGC 191
|||||:|||||
Db 32 IleGlnSerPheGluAlaLysProAspPheLeuLeuLeuLeuLeuLeuLeuLeuLeu 51
|||||:|||||
Qy 192 ACCAGCTTGTGTCAGGAGTGTCTACTTGTGTAGCCAGCGCTGACCCCGATGAGATC 251
|||||:|||||
Db 52 ThrThrTrpIleGlnGluIleValAspMetIleGluGlnAsnGlyAspValGluLysCys 71
|||||:|||||
Qy 252 GGCTTGATGAACATCGACGAGCTCCCGCTCTGGAGTAC-----CCACAGCGC 302
|||||:|||||
Db 72 GlnArgAlaIleGlnHieArgHisProPheIleGluTrpAlaArgProProGlnPro 91
|||||:|||||
Qy 303 ---GGCTTGGACATCATCAAGAACTGACCTCTCCGCGCTCATCAAGAGCCACTGCC 359
|||||:|||||
Db 92 SerGlyValGluLysAlaLysAlaMetProSerProArgIleLeuLysThrHisLeuSer 111
|||||:|||||
Qy 360 TACCGCTTCTGCGCTCTGACCTCCACAATGGAGACTCCCAAGGTCACTATATGCTCGC 419
|||||:|||||
Db 112 ThrGlnLeuLeuProProSerPheTrpGluAsnAenCysLysPheLeuTyrValAlaArg 131
|||||:|||||
Qy 420 AACCCCAAGATCTGGTGTCTTATATATCAGTTTCCACCGCTCTCTCGCGACCATGAGC 479
|||||:|||||
Db 132 AsnAlaLysAspCysMetValSerTyrTyrHisPheGlnArgMetAsnHisMetLeuPro 151
|||||:|||||
Qy 480 TACCGAGCACCTTCAAGAACTTCTCGCGAGGTTTATGATGATAGCTGGGTACGGC 539
|||||:|||||
Db 152 AspProGlyThrTrpGluLysPheGluThrPheIleAsnGlyLysValValTrpGly 171
|||||:|||||
Qy 540 TCCTGGTTGAGCAGCTGACGAGTTCCTGGGAGCACCGCATCGGACTCGAAGCTGCTTTT 599
|||||:|||||
```

```

Db      92 SerGlyValGluLysAlaLysAlaMetProSerProArgIleLeuLysThrHisLeuSer 111
Qy      360 TACCGCTTTGCGCCCTCTGAGACTCCACCAATGAGAGACTCAAGGTTCATCTATATGCTCGC 419
Db      112 ThrGlnLeuLeuProSerPheTrpGluAsnAsnCysLysPheLeuTyrValAlaArg 131
Qy      420 AACCCCAAGGATGCTGCTGCTTATATATCAGTTTCCACCGCTCTCTCGGACCAATGAGC 479
Db      132 AsnAlaLysAspCysMetValSerTyrTyrHisPheGlnArgMetAsnHisMetLeuPro 151
Qy      480 TACCGAGCCACTTTCAAGATTCTCGCGAGGTTTATGAATGATAAGCTGGCTACGGC 539
Db      152 AspProGlyThrTrpGluGlyLysPheGluThrPheIleAsnGlyLysValValTrpGly 171
Qy      540 TCCTGGTTGAGCAGCTGCAGAGTTCTGGGAGCACCGCATGGACTCGAAGCGTGTCTTTT 599
Db      172 SerTrpPheAspHisValLysGlyTrpTrpGluMetLysAspArgHisGlnIleLeuPhe 191
Qy      600 CTCAGTATGAAGACATCATCGGACCTGGTGACGATGGTGAGCAGCTGGCCAGATTC 659
Db      192 LeuPheTyrGluAspIleLysArgAspProLysHisGluIleArgLysValMetGlnPhe 211
Qy      660 CTGGGGGTCTCTGTGACAAGCCCGCTGGAGCCCTGACGGAGCAC----- 707
Db      212 MetGlyLysLysValAspGluThrValLeuAspLysIleValGlnGluThrSerPheGlu 231
Qy      708 -----TGCCACCACTGGTGACCAG 728
Db      232 LysMetLysGluAsnProMetThrAsnArgSerThrValSerLysSerIleLeuAspGln 251
Qy      729 TGCTGCAACGCTGAGCCCTCGCCGTGGCGGGGAGAGTTGGCTGTGGAAGACATC 788
Db      252 SerIleSerPhe-----MetArgLysGlyThrValGlyAspTrpLysAsnHis 268
Qy      789 TTCACCTCTCCATGAATGAGAAGTTTCAGCTGGTGTGTATTAACAGAAGATGGAAAGTGT 848
Db      269 PheThrValAlaGlnAsnGluArgPheAspGluIleTyrArgArgLysMetGluGlyThr 288
Qy      849 GACCTCACGTTT 860
Db      289 SerIleAsnPhe 292

RESULT 6
US-09-949-016-10328
; Sequence 10328, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10328
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10328

Alignment Scores:
Pred. No.: 7,52e-34 Length: 309
Score: 435.00 Matches: 96
Percent Similarity: 52.7% Conservative: 59
Best Local Similarity: 32.7% Mismatches: 121

```

```

Query Match: 9.9% Indels: 18
DB: 2 Gaps: 3
US-10-768-158-1 (1-2419) x US-09-949-016-10328 (1-309)
Qy      33 GAGCCGAGAGCCCGCCAGCACCCCGGGGAGTTCAGAGCAAGTACTTCGAGTTCATGGC 92
Db      12 GlnCysThrThrMetAsnSerGluLeuAspTyrTyrGluLysPheGluValHisGly 31
Qy      93 GTGGCGGTGCGCGCTTCTGCGCGGGGAGAGATGAGGAGATCGCAACTTCCCGGTGGG 152
Db      32 IleLeuMetTyrLysAspPheValLysTyrTrpAspAsnValGluAlaPheGlnAlaArg 51
Qy      153 CCCAGCAGCGTGTGGATGCTCCTACCTACCCCAAGTCCGGCACAGCTTCGTCGAGAGTG 212
Db      52 ProAspAspLeuValIleAlaThrTyrProLysSerGlyThrThrTrpValSerGluIle 71
Qy      213 GTCTACTTGTGTGAGCCAGGCGCTGACCCCGATGAGATCGGCTTGATGAACATCCACGAG 272
Db      72 ValTyrMetIleTyrLysGluGlyAspValGluLysCysLysGluAspValIlePheAsn 91
Qy      273 CAGCTCCCGGTCTCTGGAGTACCCACAGCG-----GGCCTGGACATCATCAAG 320
Db      92 ArgIleProPheLeuGluCysArgLysGluAsnLeuMetAsnGlyValLysGlnLeuAsp 111
Qy      321 GAACGTACCTTCCCGCCTCATCAAGACCCAGTCCGCTACCGCTTCTCCCTCTGAC 380
Db      112 GluMetAsnSerProArgIleValLysThrHisLeuProGluLeuLeuProAlaSer 131
Qy      381 CTCACAAATGAGACTCCAGGTTCATATATGCTGCAACCCCAAGGATCTGGTGTG 440
Db      132 PheTrpGluLysAspCysLysIleIleTyrLysCysArgAsnAlaLysAspValAlaVal 151
Qy      441 TCTTATTATCAGTTCCACCGCTCTCTCGGACCATGAGTACCGAGGACCTTCAAGAA 500
Db      152 SerPheTyrTyrPhePheLeuMetValAlaGlyHisProAsnProGlySerPheProGlu 171
Qy      501 TTCTCCCGGAGGTTTATGAATGATAAGCTGGCTACGGCTCCTGGTTTGACACGTGACG 560
Db      172 PheValGluLysPheMetGlnGlyValProTyrGlySerTrpTyrLysHisValLys 191
Qy      561 GAGTTCTGGGAGCACCGCATGGACTCGAACTGCTTTTCTCAAGTATGAAGACATGTCAT 620
Db      192 SerTrpTrpGluLysGlyLysSerProArgValLeuPheLeuPheTyrGluAspLeuLys 211
Qy      621 CGGACCTGTGTGACGATGGTGAGCAGCTGCCAGATTCTGTGGGGGTCTCCTGTGACAAG 680
Db      212 GluAspIleArgLysGluValIleLysLeuIleHisPheLeuGluArgLysPheSerGlu 231
Qy      681 GCCCAGCTGGAAGCCCTGACGAGCAGCTGCCAC----- 713
Db      232 GluLeuValAspArgIleIleHisThrSerPheGlnGluMetLysAsnAsnProSer 251
Qy      714 -----CAGCTGGTGACCATGCTGCAACCGCTGAGGCCCTGCC-----GTGGGC 758
Db      252 ThrAsnTyrThrThrLeuProAspGluIleMetAsnGlnLysLeuSerProPheMetArg 271
Qy      759 CGGGGAAGAGTTGGCTGTGGAAGGACATCTTCAACCTCTCCATCAATGAGAGTTTGCAC 818
Db      272 LysGlyIleThrGlyAspTrpLysAsnHisPheThrValAlaLeuAsnGluLysPheAsp 291
Qy      819 TTGGTGTATAAACAGAAAGATGGGAAAGTGTGACCTCACGTTT 860
Db      292 LysHisTyrGluGlnGlnMetLysGluSerThrLeuLysPhe 305

RESULT 7
US-08-325-562-2
; Sequence 2, Application US/08325562
; Patent No. 5714594
; GENERAL INFORMATION:
; APPLICANT: Weinsilboun, Richard M.
; APPLICANT: Aksoy, Ibrahim A.
; APPLICANT: Wood, Thomas C.

```



```
Db 227 ThrSerPheGlnGluMetLysAsnProSerThrAsnTyrThrMetMetProGluGlu 246
Qy 729 TGCTGCAACGCTGAGCCCTGCC---GTGGCCGGGGAAGAGTTGGCTGTGGAAGAC 785
Db 247 MetMetAsnGlnLysValSerProPheMetArgLysGlyIleGlyAspTrpLysAsn 266
Qy 786 ATCTTCACCGTCTCCATGAATGAGAAGTTTACATGGTGTATAAACACAGAACATGGAAAG 845
Db 267 HisPheProGluAlaLeuArgGluArgPheAspGluHisTyrLysGlnMetLysAsp 286
Qy 846 TGTGACCTCACGTTT 860
Db 287 CysThrValLysPhe 291

RESULT 10
US-09-150-141-13
; Sequence 13, Application US/09150141B
; Patent No. 6071732
; GENERAL INFORMATION:
; APPLICANT: The Board of Regents of the University of Oklahoma
; TITLE OF INVENTION: TYROSYLPROTEIN SULFOTRANSFERASES, NUCLEIC ACIDS ENCODING
; FILE REFERENCE: 5820.495
; CURRENT APPLICATION NUMBER: US/09/150.141B
; CURRENT FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: Wordperfect 5.1 (saved in ASCII format)
; SEQ ID NO 13
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: estrogen sulfotransferase
US-09-150-141-13

Alignment Scores:
Pred. No.: 2,28e-33 Length: 295
Score: 430.00 Matches: 94
Percent Similarity: 53.0% Conservative: 57
Best Local Similarity: 33.0% Mismatches: 116
Query Match: 9.8% Indels: 18
DB: 2 Gaps: 3

US-10-768-158-1 (1-2419) x US-09-150-141-13 (1-295)
Qy 60 GAGTTCGAGACAAAGTACTTCCAGTTCCATCGCGTGGCGCTGCCCGCTTCTGCGCGGG 119
Db 7 GlutTyrGluValPheGlyGluPheArgGlyValLeuMetAspLysArgPheThrLys 26
Qy 120 AAGATGGAGGAGATCGCCAACTTCCCGTGGCGCCGACGAGCTGTGGATCGTACCTAC 179
Db 27 TyrTrpGluAspValGluMetPheLeuAlaArgProAspAspLeuValIleAlaThrTyr 46
Qy 180 CCCAAGTCCCGCACCAGCTGTGTCGAGGAGTGTCTACTTGGTGAGCCAGCGGCGTCA 239
Db 47 ProlLysSerGlyThrTrpLysSerGluValValTyrMetIleTyrLysGluGlyAsp 66
Qy 240 CCCGATGAGATCGGCTTGATGAACATCCAGCAGCAGCTCCCGTCTCGAGTACCCACAG 299
Db 67 ValGluLysCysLysGluAspAlaIlePheAsnArgIleProTyrLeuGluCysArgAsn 86
Qy 300 CCG-----GGCTGGACATCATCAGNACTGACCTCTCCCGCTCATCAG 347
Db 87 GluAspLeuIleAsnGlyIleLysGlnLeuLysGluLysGluSerProArgIleValLys 106
Qy 348 AGCCACCTGCGCTACCGCTTCTGCGCTCTGACCTCCACATGAGAGATCCCAAGGTCA 407
Db 107 ThrHisLeuProLysValLeuProAlaSerPheTrpGluLysAsnCysLysMetIle 126
Qy 408 TATATGGCTCGCAACCCCAAGATCTGGTGTCTTATATATATATATATATATATATATAT 467
Db 127 TyrLeuCysArgAsnAlaLysAspValAlaValSerTyrTyrTyrPheLeuLeuMetIle 146
```

```
Qy 468 CGGACCATGAGCTACCGAGCACCTTTCAAGAATTTCTCCGAGGTTTATGATGATAAG 527
Db 147 ThrSerTyrProAsnProLysSerPheSerGluPheValGluLysPheMetGlnGlyGln 166
Qy 528 CTGGGCTACGCTCCTCGTTTGGACACGTGCAGAGTTCTGGGAGCACCCGATCGACTCG 587
Db 167 ValProTyrGlySerTrpTyrAspHisValLysAlaTrpTrpGluLysSerLysAsnSer 186
Qy 588 AACGTGCTTTTCTCAAGTATGAAGACATGATCGGACCTGGTGAGCATGTTGGAGGAC 647
Db 187 ArgValLeuPheMetPheTyrGluAspMetLysGluAspIleArgGluValValLys 206
Qy 648 CTGGCCAGATTCCTGGGGGTCTCTGTGACAAAGCCAGCTGGGAAGCCCTCACCGAGC 707
Db 207 LeuIleGluPheLeuGluArgLysProSerAlaGluLeuValAspArgIleGlnHis 226
Qy 708 TGCCAC-----CAGCTGGTGGACAC 728
Db 227 ThrSerPheGlnGluMetLysAsnProSerThrAsnTyrThrMetMetProGluGlu 246
Qy 729 TGCTGCAACGCTGAGCCCTGCC---GTGGCCGGGGAAGAGTTGGCTGTGGAAGAC 785
Db 247 MetMetAsnGlnLysValSerProPheMetArgLysGlyIleGlyAspTrpLysAsn 266
Qy 786 ATCTTCACCGTCTCCATGAATGAGAAGTTTACATGGTGTATAAACACAGAACATGGAAAG 845
Db 267 HisPheProGluAlaLeuArgGluArgPheAspGluHisTyrLysGlnMetLysAsp 286
Qy 846 TGTGACCTCACGTTT 860
Db 287 CysThrValLysPhe 291

RESULT 11
US-09-374-493-13
; Sequence 13, Application US/09374493
; Patent No. 6204016
; GENERAL INFORMATION:
; APPLICANT: The Board of Regents of the University of Oklahoma
; TITLE OF INVENTION: TYROSYLPROTEIN SULFOTRANSFERASES AND METHODS OF USE THEREOF
; FILE REFERENCE: 5820.546
; CURRENT APPLICATION NUMBER: US/09/374.493
; CURRENT FILING DATE: 1999-08-13
; EARLIER APPLICATION NUMBER: 09/150.133
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 60/072.994
; EARLIER FILING DATE: 1998-01-29
; EARLIER APPLICATION NUMBER: PCT/US99/16750
; EARLIER FILING DATE: 1999-07-23
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: Wordperfect 8.0 (saved in ASCII format)
; SEQ ID NO 13
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: estrogen sulfotransferase
US-09-374-493-13

Alignment Scores:
Pred. No.: 2,28e-33 Length: 295
Score: 430.00 Matches: 94
Percent Similarity: 53.0% Conservative: 57
Best Local Similarity: 33.0% Mismatches: 116
Query Match: 9.8% Indels: 18
DB: 2 Gaps: 3

US-10-768-158-1 (1-2419) x US-09-374-493-13 (1-295)
Qy 60 GAGTTCGAGACAAAGTACTTCCAGTTCCATCGCGTGGCGCTGCCCGCTTCTGCGCGGG 119
Db 7 GlutTyrGluValPheGlyGluPheArgGlyValLeuMetAspLysArgPheThrLys 26
```

```
Qy 120 AAGATGGAGGATCGCAACTTCCCGTGGCGCCAGCGAGTGTGGATCGTCACCTAC 179
Db 120 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 27 TyrTrpGluAspValGluMetPheLeuAlaValProAspAspLeuValIleAlaThrTyr 46
Db 27 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 180 CCCAAGTCCGACCGAGCTTCTGCGAGGAGTGTCTACTTGGTGGAGCCGCGCTGAC 239
Db 180 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 47 ProLysSerGlyThrTrpIleSerGluValValTyrMetIleTyrLysGluGlyAsp 66
Db 47 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 240 CCCGATGAGATCGGCTTGATGAACATCGACGAGCTCCCGCTCGTGGAGTACCCACAG 299
Db 240 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 67 ValGluLysCysLysGluAspAlaIlePheAsnArgIleProTyrLeuGluCysArgAsn 86
Db 67 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 300 CCG-----GGCTGGACATCATCAAGAACTGACCTCCCGCTCATCAAG 347
Db 300 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 87 GluAspLeuIleAsnGlyLysGlnLeuLysGluLysGluSerProArgIleValLys 106
Db 87 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 348 AGCCACCTGCGCTACCGGCTTCTGCGCTCCACCTCCCAATGGAGACTCCAAGTCTATC 407
Db 348 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 107 ThrHisLeuProLysValLeuProLysPheTrpGluLysAsnCysLysMetIle 126
Db 107 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 408 TATATGGCTCGCAACCCCAAGGATCTGGTGTGTCTTATTATCATGTTCCACCGCTCTCTG 467
Db 408 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 127 TyrLeuCysArgAsnAlaLysAspValAlaValSerTyrTyrPheLeuLeuMetIle 146
Db 127 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 468 CGGACCATGAGCTACCGAGGACCTTCAAGAACTTCTCGCGAGGTTTATGAATGATAAG 527
Db 468 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 147 ThrSerTyrProAsnProLysSerPheSerGluPheValGluLysPheMetGlnGlyGln 166
Db 147 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 528 CTGGCTACGCTCCCTGGTTTGAGCAGCTGAGAGTTCTGGGAGCACCGCATCGACTCG 587
Db 528 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 167 ValProTyrGlySerTrpTyrAspHisValValAlaTrpTrpGluLysSerLysAsnSer 186
Db 167 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 588 AACGTGCTTTTCTCAAGATGAACATCGATCGGAGCTGTGTGACGATGGTGGAGCAG 647
Db 588 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 187 ArgValLeuPheMetPheTyrGluAspMetLysGluAspIleArgGluValValLys 206
Db 187 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 648 CTGGCCAGATCTCTGGGGGTGCTCTGTGACAAGCCCGAGCCCTGGAAGCCCTGACGAGCAG 707
Db 648 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 207 LeuIleGluPheLeuGluArgLysProSerAlaGluLeuValAspArgIleGlnHis 226
Db 207 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 708 TGCCAC-----CAGCTGGTGGAGCAG 728
Db 708 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 227 ThrSerPheGlnGluMetLysAsnAsnProSerThrAsnTyrThrMetProGluGlu 246
Db 227 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 729 TGCTGCAACGCTGAGGCCCTGCCCC---GTGGGCGGGGAAGAGTTGGGCTGTGAAGGAC 785
Db 729 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 247 MetMetAsnGlnLysValSerProPheMetArgLysGlyIleIleGlyAspTrpLysAsn 266
Db 247 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 786 ATCTTCCGCTCTCCATGAATGAGAAGTTTGACTTGGTGTATATAACAGAGATGGGAAG 845
Db 786 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 267 HisPheProGluAlaLeuArgGluArgPheAspGluHisTyrLysGlnMetLysAsp 286
Db 267 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 846 TGTGACCTCAGGTTT 860
Db 846 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 287 CysThrValLysPhe 291
Db 287 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 12
US-09-374-824-13
; Sequence 13, Application US/09374824
; Patent No. 6207414
; GENERAL INFORMATION:
; APPLICANT: The Board of Regents of the University of Oklahoma
; TITLE OF INVENTION: TYROSYLPROTEIN SULFOTRANSFERASES AND METHODS OF USE THEREOF
; FILE REFERENCE: 5820.547
; CURRENT APPLICATION NUMBER: US/09/374,824
; CURRENT FILING DATE: 1999-08-13
; EARLIER APPLICATION NUMBER: 09/150,133
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 60/072,994
; EARLIER FILING DATE: 1998-01-29
; EARLIER APPLICATION NUMBER: PCT/US99/16750
; EARLIER FILING DATE: 1999-07-23
; NUMBER OF SEQ ID NOS: 45
```

```
; SOFTWARE: Wordperfect 8.0 (saved in ASCII format)
; SEQ ID NO 13
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: estrogen sulfotransferase
US-09-374-824-13

Alignment Scores:
Pred No.: 2,28e-33 Length: 295
Score: 430.00 Matches: 94
Percent Similarity: 53.0% Conservative: 57
Best Local Similarity: 33.0% Mismatches: 116
Query Match: 9.8% Indels: 18
DB: 2 Gaps: 3

US-10-768-158-1 (1-2419) x US-09-374-824-13 (1-295)
Qy 60 GAGTTCGAGACCAAGTACTTCGAGTTCATGGCTGGCTGCCGCCCTTCTCGCGCGG 119
Db 60 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 7 GluTyrTyrGluValPheGlyGluPheArgGlyValLeuMetAspLysArgPheThrLys 26
Db 7 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 120 AAGATGAGGAGATCGCAACTTCCCGTGGCGCCAGCGACGTGTGGATCGTCACCTAC 179
Db 120 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 27 TyrTrpGluAspValGluMetPheLeuAlaArgProAspAspLeuValIleAlaThrTyr 46
Db 27 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 180 CCCAAGTCCGCGACCGACTTGTGTCAGAGGTGTCTACTTGGTGGAGCCGCGCTGAC 239
Db 180 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 47 ProLysSerGlyThrTrpIleSerGluValValTyrMetIleTyrLysGluGlyAsp 66
Db 47 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 240 CCCGATGAGATCGGCTTGATGAACATCGACGAGCTCCCGCTCGTGGAGTACCCACAG 299
Db 240 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 67 ValGluLysCysLysGluAspAlaIlePheAsnArgIleProTyrLeuGluCysArgAsn 86
Db 67 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 300 CCG-----GGCTGGACATCATCAAGAACTGACCTCCCGCTCATCAAG 347
Db 300 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 87 GluAspLeuIleAsnGlyIleLysGlnLeuLysGluLysGluSerProArgIleValLys 106
Db 87 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 348 AGCCACCTGCGCTACCGGCTTCTGCGCTCCACCTCCCAATGGAGACTCCAAGTCTATC 407
Db 348 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 107 ThrHisLeuProLysValLeuProLysPheTrpGluLysAsnCysLysMetIle 126
Db 107 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 408 TATATGGCTCGCAACCCCAAGGATCTGGTGTGTCTTATTATCATGTTCCACCGCTCTCTG 467
Db 408 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 127 TyrLeuCysArgAsnAlaLysAspValAlaValSerTyrTyrPheLeuLeuMetIle 146
Db 127 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 468 CGGACCATGAGCTACCGAGGACCTTCAAGAACTTCTCGCGAGGTTTATGAATGATAAG 527
Db 468 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 147 ThrSerTyrProAsnProLysSerPheSerGluPheValGluLysPheMetGlnGlyGln 166
Db 147 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 528 CTGGCTACGCTCCCTGGTTTGAGCAGCTGAGAGTTCTGGGAGCACCGCATCGACTCG 587
Db 528 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 167 ValProTyrGlySerTrpTyrAspHisValValAlaTrpTrpGluLysSerLysAsnSer 186
Db 167 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 588 AACGTGCTTTTCTCAAGATGAACATCGATCGGAGCTGTGTGACGATGGTGGAGCAG 647
Db 588 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 187 ArgValLeuPheMetPheTyrGluAspMetLysGluAspIleArgGluValValLys 206
Db 187 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 648 CTGGCCAGATCTCTGGGGGTGCTCTGTGACAAGCCCGAGCCCTGGAAGCCCTGACGAGCAG 707
Db 648 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 207 LeuIleGluPheLeuGluArgLysProSerAlaGluLeuValAspArgIleGlnHis 226
Db 207 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 708 TGCCAC-----CAGCTGGTGGAGCAG 728
Db 708 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 227 ThrSerPheGlnGluMetLysAsnAsnProSerThrAsnTyrThrMetProGluGlu 246
Db 227 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 729 TGCTGCAACGCTGAGGCCCTGCCCC---GTGGGCGGGGAAGAGTTGGGCTGTGAAGGAC 785
Db 729 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 247 MetMetAsnGlnLysValSerProPheMetArgLysGlyIleIleGlyAspTrpLysAsn 266
Db 247 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 786 ATCTTCCGCTCTCCATGAATGAGAAGTTTGACTTGGTGTATATAACAGAGATGGGAAG 845
Db 786 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 267 HisPheProGluAlaLeuArgGluArgPheAspGluHisTyrLysGlnMetLysAsp 286
Db 267 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 846 TGTGACCTCAGGTTT 860
Db 846 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 287 CysThrValLysPhe 291
Db 287 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Db 267 HisPheProGluAlaLeuArgGluArgPheAspGluHisTyrLysGlnGlnMetLysAsp 286
Qy 846 TGTGACCTCAGTTT 860
||| :|||
Db 287 CysThrValLysPhe 291
||| :|||
RESULT 13
US-09-374-492-13
; Sequence 13, Application US/09374492
; Patent No. 6207432
; GENERAL INFORMATION:
; APPLICANT: The Board of Regents of the University of Oklahoma
; TITLE OF INVENTION: TYROSYLPROTEIN SULFOTRANSFERASES AND METHODS OF USE THEREOF
; FILE REFERENCE: 5820.545
; CURRENT APPLICATION NUMBER: US/09/374,492
; CURRENT FILING DATE: 1999-08-13
; EARLIER APPLICATION NUMBER: 09/150,141
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 60/072,994
; EARLIER FILING DATE: 1998-01-29
; EARLIER APPLICATION NUMBER: PCT/US99/16750
; EARLIER FILING DATE: 1999-07-23
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: Wordperfect 8.0 (saved in ASCII format)
; SEQ ID NO 13
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: estrogen sulfotransferase
US-09-374-492-13
Alignment Scores:
Pred. No.: 2,28e-33 Length: 295
Score: 430.00 Matches: 94
Percent Similarity: 53.0% Conservative: 57
Best Local Similarity: 33.0% Mismatches: 116
Query Match: 9.8% Indels: 18
Gaps: 3
DB:
US-10-768-158-1 (1-2419) x US-09-374-492-13 (1-295)
Qy 60 GAGTTCGAGCAAGTACTTCAGTTCATCGCGTGGCGTCCGCCCTTCGCGCGG 119
||| :|||
Db 7 GluTyrGluValPheGlyGluPheArgGlyValLeuMetAspLysArgPheThrLys 26
||| :|||
Qy 120 AAGATGGAGGAGATCGCAACTTCCCGTGGCGGCCCGAGCGAGTGTGGATCGTCACCTAC 179
||| :|||
Db 27 TyrTrpGluAspValGluMetPheLeuAlaArgProAspLeuValIleAlaThrTyr 46
||| :|||
Qy 180 CCCAAGTCCGCGACAGCTTCTCGCAGGAGTGTCTACTTGTGGTGGAGCCGCGCTGAC 239
||| :|||
Db 47 ProlYsSerGlyThrThrTrpIleSerGluValValTyrMetIleTyrLysGluLysAsp 66
||| :|||
Qy 240 CCCGATGATCGGCTTGTATGAACATCGACGAGCGTCCCGTCCGCTCGAGTACCCACAG 299
||| :|||
Db 67 ValGluLysCysLysGluAspAlaIlePheAsnArgIleProTyrLeuGluCysArgAsn 86
||| :|||
Qy 300 CCG-----GGCTCGACATCATCAAGGAAGTACCTCTCCCGCGCTCATCAAG 347
||| :|||
Db 87 GluAspLeuIleAsnGlyIleLysGlnLeuLysGluLysGluSerProArgIleValLys 106
||| :|||
Qy 348 AGCCACCTGCGCTACCGCTTCTGCGCTCTGACCTCCCAATGGAGACTCCAGGTCTATC 407
||| :|||
Db 107 ThrHisLeuProLysValLeuProAlaSerPheTrpGluLysAsnCysLysMetIle 126
||| :|||
Qy 408 TATATGGCTCGCAACCCCAAGGATCTGGTGTGTCTTATTATCATGTTCCACCGCTCTCTG 467
||| :|||
Db 127 TyrLeuCysArgAsnAlaLysAspValAlaValSerTyrTyrPheLeuLeuMetIle 146
||| :|||
Qy 468 CGGACCATGAGTACCGGAGGACCTTTCAGAATTCGCGGAGGTTTATGAATGATAG 527
||| :|||
Db 147 ThrSerTyrProAsnProLysSerPheSerGluPheValGluLysPheMetGlnGlyGln 166
||| :|||

Qy 528 CTGGGCTACGGCTCCTCGTTTGGACACGTGCAGAGTTCGGAGACCCGCGTGGACTCG 587
||| :|||
Db 167 ValProTyrGlySerTrpTyrAspHisValLysAlaTrpTrpGluLysSerLysAsnSer 186
||| :|||
Qy 588 AACGTGCTTTTCTCAAGTATGAAGACATGATCGGAGCCTGGTGGTGCATGCTGGAGCAG 647
||| :|||
Db 187 ArgValLeuPheMetPheTyrGluAspMetLysGluAspIleArgArgGluValValLys 206
||| :|||
Qy 648 CTGGCCAGATTCTCGGGGTGCTCTGTGACAAGCCCGAGCTGGAGCCCTGACGAGCAG 707
||| :|||
Db 207 LeuIleGluPheLeuGluArgLysProSerAlaGluLeuValAspArgIleIleGlnHis 226
||| :|||
Qy 708 TGCAC-----CAGCTGGTGGACAG 728
||| :|||
Db 227 ThrSerPheGlnGluMetLysAsnAsnProSerThrAsnTyrThrMetMetProGluGlu 246
||| :|||
Qy 729 TGTGCAACGCTGAGGCCCTGCC---GTGGCGCGGGAAGAGTTGGGCTGTGGAAGAC 785
||| :|||
Db 247 MetMetAsnGlnLysValSerProPheMetArgLysGlyIleIleGlyAspTrpLysAsn 266
||| :|||
Qy 786 ATCTTACCGTCTCCATGAATGAGAGTTCGACTTGTGTATATAACAGAGATGGGAAG 845
||| :|||
Db 267 HisPheProGluAlaLeuArgGluArgPheAspGluHisTyrLysGlnGlnMetLysAsp 286
||| :|||
Qy 846 TGTGACCTCAGTTT 860
||| :|||
Db 287 CysThrValLysPhe 291
||| :|||
RESULT 14
US-09-785-343-13
; Sequence 13, Application US/09785343
; Patent No. 6605455
; GENERAL INFORMATION:
; APPLICANT: The Board of Regents of the University of Oklahoma
; TITLE OF INVENTION: TYROSYLPROTEIN SULFOTRANSFERASES AND METHODS OF USE THEREOF
; FILE REFERENCE: 5827.003
; CURRENT APPLICATION NUMBER: US/09/785,343
; CURRENT FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: 09/150,133
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/072,994
; PRIOR FILING DATE: 1998-01-29
; PRIOR APPLICATION NUMBER: PCT/US99/16750
; PRIOR FILING DATE: 1999-07-23
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: Wordperfect 8.0 (saved in ASCII format)
; SEQ ID NO 13
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: estrogen sulfotransferase
US-09-785-343-13
Alignment Scores:
Pred. No.: 2,28e-33 Length: 295
Score: 430.00 Matches: 94
Percent Similarity: 53.0% Conservative: 57
Best Local Similarity: 33.0% Mismatches: 116
Query Match: 9.8% Indels: 18
Gaps: 3
DB:
US-10-768-158-1 (1-2419) x US-09-785-343-13 (1-295)
Qy 60 GAGTTCGAGCAAGTACTTCAGTTCATCGCGTGGCGTCCGCCCTTCGCGCGG 119
||| :|||
Db 7 GluTyrGluValPheGlyGluPheArgGlyValLeuMetAspLysArgPheThrLys 26
||| :|||
Qy 120 AAGATGGAGGAGATCGCAACTTCCCGTGGCGGCCCGAGCGAGTGTGGATCGTCACCTAC 179
||| :|||
Db 27 TyrTrpGluAspValGluMetPheLeuAlaArgProAspLeuValIleAlaThrTyr 46
||| :|||

Db 81 AlaGlnThrLeuAspArgHisAlaPheLeuLeuLeuLysPheProHisLysLeuLysPro 100
Qy 303 GGCCTGGACATCATCAAGAACTGACCTCTCCCGCCCTCATCAAGAGCACCTGCCCTAC 362
Db 101 AspLeuGluPheValLeuLeuMetSerSerProGlnLeuLeuLysThrHisLeuProSer 120
Qy 363 GCCTTTCTGCCCTCTGACCTCCACATCGAGACTCAAGGTCATATATATATGCTGCAAC 422
Db 121 HisLeuIleProProSerIleTrpLysGluAsnCysLysIleValTyValAlaArgAsn 140
Qy 423 CCCAAGGATCTGGTCTCTAT 482
Db 141 ProllysAspCysLeuValSerIleHisPheHisArgMetAlaSerPheMetProAsp 160
Qy 483 CGAGGCACCTTTCAGAAATTCGCGGAGGTTTATGAATGATGAAGCTGGGCTACGGCTCC 542
Db 161 ProGlnAsnLeuGluPheTyTrpLysPheMetSerGlyLysValValGlySer 180
Qy 543 TGGTTTGAGCAGTGCAGGAGTTCCTGGAGCACCGCATGGATCGAAGCTGCTTTTCTC 602
Db 181 TrpPheAspHisValLysGlyTrpTrpAlaAlaLysAspMetHisArgIleLeuTyLeu 200
Qy 603 AAGTATGAAGACATCATCGGAGCTGTGGATGGTGGAGCAGCTGCCAGATTCCTG 662
Db 201 PheTyLeuAspIleLysLysAspProLysArgGluIleGluLysIleLeuLysPheLeu 220
Qy 663 GGGGTGTCTGTGACAAAGCCAGCTGGAAGCCCTGACGAGCAGCTGCCACAGCTGGT 722
Db 221 GluLysAspIleSerGluGluLeuLeuAsnLysIleIleTyHisThrSerPheAspVal 240
Qy 723 -----GACCAGTGTGCAAGCTGAGGCCCTGCGCGTGGG----- 758
Db 241 MetLysGlnAsnProMetThrAsnTyThrThrLeuProThrSerIleMetAspHisSer 260
Qy 759 -----CGGGAGAGTTCGGCTGGAGGACATCTTCACCGTCTCC 800
Db 261 IleSerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyPheThrValAla 280
Qy 801 ATGAATGAGAAGTTGACTTGGTGTATATAAACAAGAGATGGAAAGTGTGACCTCACGTTT 860
Db 281 GlnAsnGluGluPheAspLysAspTyTrpGlnLysLysMetAlaGlySerThrLeuThrPhe 300

RESULT 18

US-10-199-334-5
; Sequence 5, Application US/10199334
; Patent No. 6905855
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-3
; CURRENT APPLICATION NUMBER: US/10/199,334
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-334-5

Alignment Scores: 1.43e-32 Length: 304
Pred. No.: 422.00 Matches: 97
Score:

Percent Similarity: 52.0% Conservative: 59
Best Local Similarity: 32.3% Mismatches: 124
Query Match: 9.6% Indels: 20
DB: 2 Gaps: 5
US-10-768-158-1 (1-2419) x US-10-199-334-5 (1-304)
Qy 21 ATGCGGAGACGCGGCGGAGACCCAGCACCCGCGGGGAGTTTCGAG---AGCAAGTAC 77
Db 1 MetAlaLysIleGluLysAsnAlaProThrMetGluLysLysProGluLeuPheAsnIle 20
Qy 78 TTCAGATTCCATGCGGTGCGGCTCCGCCCTTCTGCGCGGGAAGATGAGGAGATCGCC 137
Db 21 MetGluValAspGlyValProThrLeuLeuSerLysGluTrpTrpGluLysValCys 40
Qy 138 AACTTCCCGGTGCGGCGGCGGAGCTGTGGATGTCTACCTACCCCAAGTCCGGCACCAGC 197
Db 41 AsnPheGlnAlaLysProAspAspLeuLeuAlaThrTyProLysSerGlyThrThr 60
Qy 198 TTGCTGAGGAGTGTCTACTTGTGTGAGCGGCGCTGACCCCGATGAGATCGGCTTG 257
Db 61 TrpMetHisGluIleLeuAspMetIleLeuAsnAspGlyAspValGluLysCysLysArg 80
Qy 258 ATGAACATCGACGAGAGCTCCCGTCTGGAG-----TACCCA-----CAGCG 302
Db 81 AlaGlnThrLeuAspArgHisAlaPheLeuGluLysPheProHisLysLysGluLysPro 100
Qy 303 GGCCTGGACATCATCAAGAACTGACCTCTCCCGCCCTCATCAAGAGCACCTGCCCTAC 362
Db 101 AspLeuGluPheValLeuLeuMetSerSerProGlnLeuLysIleTyHisThrSer 120
Qy 363 CGCTTTCTGCCCTCTGACCTCCACAAATGGAGACTCCAAAGTTCATATATATGCTCGCAAC 422
Db 121 HisLeuIleProProSerIleTrpLysGluAsnCysLysIleValTyValAlaArgAsn 140
Qy 423 CCCAAGATCTGGTGTCTTAT 482
Db 141 ProllysAspCysLeuValSerTyTrpHisPheHisArgMetAlaSerPheMetProAsp 160
Qy 483 CGAGGCACCTTTCAGAAATTCGCGGAGGTTTATGAATGATGAAGTGTGGCTACGGCTCC 542
Db 161 ProGlnAsnLeuGluPheTyTrpLysPheMetSerGlyLysValValGlySer 180
Qy 543 TGGTTTGAGCAGTGCAGGAGTTCCTGGAGCACCGCATGGATCGAAGCTGCTTTTCTC 602
Db 181 TrpPheAspHisValLysGlyTrpTrpAlaAlaLysAspMetHisArgIleLeuTyLeu 200
Qy 603 AAGTATGAAGACATCATCGGAGCTGTGGATGGTGGAGCAGCTGCCAGATTCCTG 662
Db 201 PheTyLeuAspIleLysLysAspProLysArgGluIleGluLysIleLeuLysPheLeu 220
Qy 663 GGGGTGTCTGTGACAAAGCCAGCTGGAAGCCCTGACGAGCAGCTGCCACAGCTGGT 722
Db 221 GluLysAspLysSerGluGluIleLeuAsnLysIleIleTyHisThrSerPheAspVal 240
Qy 723 -----GACCAGTGTGCAAGCTGAGGCCCTGCGCGTGGG----- 758
Db 241 MetLysGlnAsnProMetThrAsnTyThrThrLeuProThrSerIleMetAspHisSer 260
Qy 759 -----CGGGAGAGTTCGGCTGGAGGACATCTTCACCGTCTCC 800
Db 261 IleSerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyPheThrValAla 280
Qy 801 ATGAATGAGAAGTTTGAATGCTGTATATAAACAAGAGATGGAAAGTGTGACCTCACGTTT 860
Db 281 GlnAsnGluGluPheAspLysAspTyTrpGlnLysLysMetAlaGlySerThrLeuThrPhe 300

RESULT 19

US-10-199-329-5
; Sequence 5, Application US/10199329
; Patent No. 6953681
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.

```
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-2
; CURRENT APPLICATION NUMBER: US/10/199,329
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-199-329-5

Alignment Scores:
Pred. No.: 1.43e-32 Length: 304
Score: 422.00 Matches: 97
Percent Similarity: 52.0% Conservative: 59
Best Local Similarity: 32.3% Mismatches: 124
Query Match: 9.6% Indels: 20
DB: 2 Gaps: 5

US-10-768-158-1 (1-2419) x US-10-199-329-5 (1-304)
QY 21 ATGCGGAGAGCGGAGCCGAGACCCCGGAGGAGTTCGAG---AGCAAGTAC 77
DB 1 MetAlaLysIleGluLeuAsnAlaProThrMetGluLysLysProGluLeuPheAsnIle 20
QY 78 TTCGAGTTCATCGCTGCGGCTGCGGCTTCGCGCGGAGAGTGGAGAGATCGCC 137
DB 21 MetGluValaspGlyValProThrLeuLeuSerLysGluTrpTrpGluLysValCys 40
QY 138 AACTTCCGGTCCGCGGAGGAGTGTGGTACCTACCCCAAGTCCGCGGAGCAGC 197
DB 41 AsnPheGlnAlaLysProAspAspLeuLeuAlaThrTyrProLysSerGlyThrThr 60
QY 198 TTGCTGAGGAGGTGCTTACTTGGTACCGGCGCTGACCCCGATGAGATCGGCTG 257
DB 61 TrpMetHisGluIleLeuAspMetIleLeuAsnAspGlyAspValGluLysCysLysArg 80
QY 258 ATGAACATCGAGCAGCAGCTCCGCTCTCGGAG-----TACCCA-----CAGCG 302
DB 81 AlaGlnThrLeuAspArgHisAlaPheLeuGluLeuLysPheProHisLysGluLysPro 100
QY 303 GGCCTGGACATCATCAAGGAATGACCTCTCCCGGCTCATCAAGAGCCACTGCCCTAC 362
DB 101 AspLeuGluPheValLeuGluMetSerSerProGlnLeuIleLysThrHisLeuProSer 120
QY 363 CGCTTTCGCTCTGACCTCCCAATCGAGACTCCAGGTCACTATATGCTCGGAC 422
DB 121 HisLeuIleProProSerIleTrpLysGluAsnCysLysIleValTyrValAlaArgAsn 140
QY 423 CCCAAGGATCGGTGCTTATATATGATTCACCGCTCTCTCGGACCATGAGTAC 482
DB 141 ProLysAspCysLeuValSerTyrHisPheHisArgMetAlaSerPheMetProAsp 160
QY 483 CGAGGCACCTTTCAGAAATTCGCGGAGGTTTATGAATGATTAAGTGGGTTCGCTCC 542
DB 161 ProGlnAsnLeuGluGluPheTyrGluLysPheMetSerGlyLysValValGlySer 180
QY 543 TGGTTTACGACCTGACGAGTTCGAGGACCCCATCGGACTCGAAAGTCTTTTCTC 602
DB 181 TrpPheAspHisValLysGlyTrpAlaLysAspMetHisArgIleLeuTyrLeu 200
QY 603 AAGTATGAAGACATGATCGGAGCCTGCTGACGATGGTGGAGCAGCTGGCCAGTTCCTG 662
```

```
DB 201 PheTyrGluAspIleLysLysAspProLysArgGluIleGluLysIleLeuLysPheLeu 220
QY 663 GGGGTGCTCTGTGACAGGCCAGCTGGAAGCCCTGACGAGCACTCCACCACTGCTG 722
DB 221 GluLysAspIleSerGluGluIleLeuAsnLysIleTyrHisThrSerPheAspVal 240
QY 723 -----CACCAGTGTGCAACGCTGAGGCCCTGCCCTGGGC----- 758
DB 241 MetLysGlnAsnProMetThrAsnTyrThrLeuProThrSerIleMetAspHisSer 260
QY 759 -----CGGGAAGAGTTGGGCTGTGGAAGGACATCTTCACCGTCTCC 800
DB 261 IleSerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAla 280
QY 801 ATGAATCAGAAAGTTGACTGTGGTGTATAAACAGAGAATGGAAAGTGTGACCTCACGTT 860
DB 281 GlnAsnGluGluPheAspLysAspTyrGlnLysLysMetAlaGlySerThrLeuThrPhe 300

RESULT 20
US-10-199-330-5
; Sequence 5, Application US/10199330
; Patent No. 6967094
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-1
; CURRENT APPLICATION NUMBER: US/10/199,330
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-199-330-5

Alignment Scores:
Pred. No.: 1.43e-32 Length: 304
Score: 422.00 Matches: 97
Percent Similarity: 52.0% Conservative: 59
Best Local Similarity: 32.3% Mismatches: 124
Query Match: 9.6% Indels: 20
DB: 2 Gaps: 5

US-10-768-158-1 (1-2419) x US-10-199-330-5 (1-304)
```

```
QY 21 ATGCGGAGAGCGGAGCGGAGACCCCGGAGGAGTTCGAG---AGCAAGTAC 77
DB 1 MetAlaLysIleGluLysAsnAlaProThrMetGluLysLysProGluLeuPheAsnIle 20
QY 78 TTCGAGTTCATCGCTGCGGCTGCGGCTTCGCGCGGAGAGATGAGGAGATCGCC 137
DB 21 MetGluValaspGlyValProThrLeuLeuSerLysGluTrpTrpGluLysValCys 40
QY 138 AACTTCCGGTCCGCGGAGGAGTGTGGATGCTACCTACCCCAAGTCCGCGGAGCAGC 197
DB 41 AsnPheGlnAlaLysProAspAspLeuLeuAlaThrTyrProLysSerGlyThrThr 60
QY 198 TTGCTGAGGAGGTGCTTACTTGGTACCGGCGCTGACCCCGATGAGATCGGCTG 257
DB 61 TrpMetHisGluIleLeuAspMetIleLeuAsnAspGlyAspValGluLysCysLysArg 80
```



```
QY 258 ATGAACATCGAGCAGCAGCTCCCGCTCTGGAG-----TACCCA-----CAGCCG 302
Db      :::::  |||||  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 303 GGCCTGGACATCATCAAGAACTACCTCCCGCTCATCAAGAGCCAGCTGCCCTAC 362
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 101 AspLeuGluPheValLeuGluMetSerProGlnLeuIleLysThrHisLeuProSer 120
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 363 CGCTTCTCTGCTGACCTCCACAATGGAGACTCCAAGGTCATCTATATGGCTCGCAAC 422
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 121 HisLeuIleProProSerIleTrpLysGluAsnCysLysIleValTyValAlaArgAsn 140
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 423 CCAAGATCTGGTGTCTTATATCAGTTCCACCGCTCTCTCGGACCATGAGCTAC 482
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 141 ProLysAspCysLeuValSerTyTrpHisPheHisArgMetAlaSerPheMetProAsp 160
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 483 CGAGGCACCTTTCAAGAACTTCGGGAGGTTTATGAATGATAGCTAGCTGGCTAGCGCTCC 542
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 161 ProGlnAsnLeuGluGluPheTyTrpLysPheMetSerGlyLysValValGlyLys 180
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 543 TGGTTTGAGCAGCTGCAGAGTTCTGGGAGCACCAGCATGGACTCGAAAGTGTCTTTTCTC 602
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 181 TrpPheAspHisValLysGlyTrpTrpAlaLysAspMetHisArgIleLeuTyTrp 200
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 603 AAGTATGAAGACATCATCGGACCTGTGTGACGATGGTGGAGCAGCTGGCCAGATTCCTG 662
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 201 PheTyTrpGluAspIleLysLysAspProLysArgGluIleGluLysIleLeuLysPheLeu 220
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 663 GGGGTGCTCTGTGACAAGCCAGCTGGAAGCCCTGACGGAGCAGCTGCCACCGCTGTG 722
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 221 GluLysAspIleSerGluGluIleLeuAsnLysIleIleTyHisThrSerPheAspVal 240
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 723 -----GACCAGTGTGCAAGCTGAGGCCCTGCCCGTGGG----- 758
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 241 MetLysGlnAsnProMetThrAsnTyTrpThrLeuProThrSerIleMetAspHisSer 260
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 759 -----CGGGGAAGATTGGGTGTGGAGGACATCTTACCGTCTCC 800
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 261 IleSerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyTrpThrValAla 280
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 801 ATGAATGAGAAGTTGACTTGGTGATATAACAGAGATGGGAAAGTGAGCTCACGTTT 860
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 281 GlnAsnGluGluPheAspLysAspTyTrpGlnLysLysMetAlaGlySerThrLeuThrPhe 300
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::

RESULT 21
US-09-609-816-8
; Sequence 8, Application US/09609816
; Patent No. 6436684
; GENERAL INFORMATION:
; APPLICANT: Woodage, Trevor
; APPLICANT: Wei, Minh Hui
; APPLICANT: Kodira, Chinappa
; APPLICANT: Beasley, Ellen
; APPLICANT: Difrancesco, Valentina
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669PCT
; CURRENT APPLICATION NUMBER: US/09/609,816
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 283
; TYPE: PRT
; ORGANISM: HUMAN
; APPLICANT: Woodage, Trevor
```

```
Alignment Scores:
Pred. No.: 2,42e-32 Length: 283
Score: 419.50 Matches: 91
Percent Similarity: 53.0% Conservative: 57
Best Local Similarity: 32.8% Mismatches: 112
Query Match: 9.5% Indels: 19
DB: Gaps: 4

US-10-768-158-1 (1-2419) x US-09-609-816-8 (1-283)

QY 81 GAGTTCATCGCTGCGCTGCCCTCTTCCCGCGGAAGATGGAGGAGATGCCCAAC 140
Db 1 GluValAsnGlyIleLeuMetSerLysLeuMetSerAspAsnTrpAspLysIleTrpAsn 20
QY 141 TTCCCGGTGCGCGCCAGCAGCTGTGTGATCTACCTACCCCAAGTCCGGCCAGCAGCTTG 200
Db 21 PheGlnAlaLysProAspAspLeuLeuIleAlaThrTyAlaLysAlaGlyThrTrp 40
QY 201 CTGACGAGGTGTCTACTTGGTGAGCCAGCGGCTGACCCCGATGAGATCGGCTTTGATG 260
Db 41 ThrGlnGluIleValAspMetIleGlnAsnAspGlyAspValGlnLysCysGlnArgAla 60
QY 261 AACATCCAGCAGCAGCTCCCGCTCTCGAGTAC-----CCACAGCCG-----GGC 305
Db 61 AsnThrTyTrpAspArgHisProPheIleGluTrpThrLeuProSerProLeuAsnSerGly 80
QY 306 CTGACACATCAAGGAAGTACCTCTCCCGCTCAAGAGCCAGCTGCCCTTACCGC 365
Db 81 LeuAspLeuAlaAsnLysMetProSerProArgThrLeuLysThrHisLeuProValHis 100
QY 366 TTCTGCGCTCTGACCTCCACAATGGAGACTCCAAAGTTCATATATGGTTCGCAACCCC 425
Db 101 MetLeuProProSerPheTrpLysGluAsnSerLysIleIleTyValAlaArgAsnAla 120
QY 426 AAGGATCTGTGGTGTCTTATTATCAGTTCACCGCTCTCTCGGACCATGAGTACCGA 485
Db 121 LysAspCysLeuValSerTyTrpPheSerArgMetAsnLysMetLeuProAspPro 140
QY 486 GGCACTTTCAGAAATCTGCCGGAGGTTTATGAATGATAGCTGGGTGAGCTGCTCCG 545
Db 141 GlyThrLeuGlyGluTyTrpIleGluGlnPheLysAlaGlyLysValLeuTrpGlySerTrp 160
QY 546 TTGAGCAGCTGCGAGGAGTTCTGGGAGCACCAGCATGGACTCGAACGTCGCTTTTCTCAAG 605
Db 161 TyrAspHisValLysGlyTrpTrpAspValLysAspGlnHisArgIleLeuTyLeuPhe 180
QY 606 TATGAACACATGATCGGACCTCGGACCTGTCAGATGTGTGGAGCAGCTGGCCAGATTCTTGGGG 665
Db 181 TyrGluAspMetLysGluAspProLysArgGluIleLysLysIleAlaLysPheLeuGlu 200
QY 666 GTGTCTGTGACACAGGCCAGCTGGAAGCCCTGACGGAGCAGCTGCCACCGCTGTG-- 722
Db 201 LysAspLysSerGluGluValLeuAsnLysIleIleTyHisThrSerPheAspValMet 220
QY 723 -----GACCAGTGTCTGCAAGCTGAGGCCCTGCCG----- 752
Db 221 LysGluAsnProMetAlaAsnTyTrpThrLeuProSerSerIleMetAspHisSerIle 240
QY 753 -----GTGGCCCGGGGAAGAGTTGGGTGTGGGAAGGACATCTTCACCGTCTCCATG 803
Db 241 SerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyTrpPheValAlaGln 260
QY 804 AATGAGAAAGTTTACTTGGTGATATAACAGAAAGATGGAAAGTGTGACCTCACGTTT 860
Db 261 SerGluAspPheAspGluAspTyTrpArgArgLysMetAlaGlySerAsnIleThrPhe 279

RESULT 22
US-09-609-816-9
; Sequence 9, Application US/09609816
; Patent No. 6436684
; GENERAL INFORMATION:
; APPLICANT: Woodage, Trevor
```


Db 81 LeuAspLeuAlaAsnLysMetProSerProArgThrLeuLysThrHisLeuProValHis 100
Qy 366 TTTCGCCCTCTGACCTCCACAAATGGAGACATCCCAAGGTCACTATATGCTCGCAACCCC 425
Db 101 MetLeuProProSerPheTrpLysGluAsnSerLysIleIleTyrValAlaArgAsnAla 120
Qy 426 AAGATCTGGTGGTCTTATATATCAGTTCCACCGCTCTCTCGGACCATCAGTACCGA 485
Db 121 LysAspCysLeuValSerTyrTyrPheSerArgMetAsnLysMetLeuProAspPro 140
Qy 486 GGCACCTTCAAGAAATCTCGCGAGGTTTATGAATGATAAGCTGGGCTACGGCTCCCTGG 545
Db 141 GlyThrLeuGlyGluTyrIleGluInPheLysAlaGlyLysValLeuTrpGlySerTrp 160
Qy 546 TTTCGACGACGTCAGGAGTTCTGGAGCACCACCGCATGGACTCGAAGCTGCTTTTCTCAAG 605
Db 161 TyrAspHisValLysGlyTrpAspValLysAspGlnHisArgIleLeuTyrLeuPhe 180
Qy 606 TATGAACATCATCGGACCTGGTACAGTGTGGAGCAGCTGGGACGCTGGCCAGATTCTCTGGG 665
Db 181 TyrGluAspMetLysGluAspProLysArgGluIleLysLysIleAlaLysPheLeuGlu 200
Qy 666 GTGTCTGTGACAAAGGCCACCTGGAAGCCCTGACGGAGCACTGCCACCACTGGTG-- 722
Db 201 LysAspIleSerGluGluValLeuAsnLysIleIleTyrHisThrSerPheAspValMet 220
Qy 723 -----GACCAGTGTGCAACGCTGAGGCCCTGCCCC----- 752
Db 221 LysGluAsnProMetAlaAsnTyrThrThrLeuProSerSerIleMetAspHisSerIle 240
Qy 753 -----GTGGCCGGGGAAGAGTTGGGCTGTGGAAGGACATCTTCACCGTCTCCATG 803
Db 241 SerProPheMetAlaLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAlaGln 260
Qy 804 AATGAGAAGTTTGACTGGTGTATAAACAGAGATGGAAAGTGTGACCTCACGTTT 860
Db 261 SerGluAspPheAspGluAspTyrArgArgLysMetAlaGlySerAsnIleThrPhe 279

RESULT 24
US-10-768-334-9
; Sequence 9, Application US/10199334
; Patent No. 6905855
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-3
; CURRENT APPLICATION NUMBER: US/10/199,334
; PRIOR FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-334-9

Alignment Scores:
Pred. No.: 2,42e-32 Length: 283
Score: 419.50 Matches: 91
Percent Similarity: 53.0% Conservative: 57
Best Local Similarity: 32.8% Mismatches: 112
Query Match: 9.5% Indels: 19
DB: 2 Gaps: 4

US-10-768-158-1 (1-2419) x US-10-199-334-9 (1-283)
Qy 81 GAGTTTCATCGCGCTCGCGCTCCGCCCTTCTGCGCGGGAAGATGGAGGATCGCCCAAC 140
Db 1 GluValAsnGlyIleLeuMetSerLysLeuMetSerAspAsnTrpAspLysIleTrpAsn 20
Qy 141 TTCCCGGTGGCGGCCAGCGAGCTGTGGATCTCCTACCTACCCCAAGTCGCGGACCAAGCTTG 200
Db 21 PheGlnAlaLysProAspAspLeuIleAlaThrTyrAlaLysAlaGlyThrThrTrp 40
Qy 201 CTGCAGAGGTGGTCTACTTGGTGAGCCAGCGGCTGACCCCGATGAGATCGGCTTGATG 260
Db 41 ThrGlnGluIleValAspMetIleGlnAsnAspGlyAspValGlnLysCysGlnArgAla 60
Qy 261 AACATCGACGACGAGCTCCCGTCTCTGAGTAC-----CCACAGCCG-----GCC 305
Db 61 AsnThrTyrAspArgHisProPheIleGluTrpThrLeuProSerProLeuAsnSerGly 80
Qy 306 CTGCACATCATCAAGAACTGACCTCTCCCGCTCATCAAGAGCCACTCCTCCCTACCGC 365
Db 81 LeuAspLeuAlaAsnLysMetProSerProArgThrLeuLysThrHisLeuProValHis 100
Qy 366 TTTCGCGCTGTGACCTCCACAAATGGAGACTCCCAAGGTCACTATATGCTCGCAACCCC 425
Db 101 MetLeuProProSerPheTrpLysGluAsnSerLysIleIleTyrValAlaArgAsnAla 120
Qy 426 AAGATCTGGTGGTGTCTTATTATCAGTTCCACCGCTCTCTGCGGACCATCAGCTACCGA 485
Db 121 LysAspCysLeuValSerTyrTyrPheSerArgMetAsnLysMetLeuProAspPro 140
Qy 486 GGCACTTTCAGAAATTTCTCGCGAGGTTTATGAATGATAAGCTGGGCTACGGCTCTCTGG 545
Db 141 GlyThrLeuGlyGluTyrIleGluInPheLysAlaGlyLysValLeuTrpGlySerTrp 160
Qy 546 TTTCGACGACGTCAGGAGTTCTGGGACACCGCATGGACTCGAAGCTGCTTTTCTCAAG 605
Db 161 TyrAspHisValLysGlyTrpTrpAspValLysAspGlnHisArgIleLeuTyrLeuPhe 180
Qy 606 TATGAACATGTCATCGGACCTGGTACGATGTGGAGCAGCTGGCGGACGATTTCTCTGGG 665
Db 181 TyrGluAspMetLysGluAspProLysArgGluIleLysLysIleAlaLysPheLeuGlu 200
Qy 666 GTGTCTGTGACAAAGGCCAGCTGGAAGCCCTGACGGAGCACTGCCACCACTGGTG--- 722
Db 201 LysAspIleSerGluGluValLeuAsnLysIleIleTyrHisThrSerPheAspValMet 220
Qy 723 -----GACCAGTGTGCAACGCTGAGGCCCTGCCCC----- 752
Db 221 LysGluAsnProMetAlaAsnTyrThrThrLeuProSerSerIleMetAspHisSerIle 240
Qy 753 -----GTGGCCGGGGAAGAGTTGGGCTGTGGAAGGACATCTTCACCGTCTCCATG 803
Db 241 SerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAlaGln 260
Qy 804 AATGAGAAGTTTGACTGGTGTATAAACAGAGATGGAAAGTGTGACCTCACGTTT 860
Db 261 SerGluAspPheAspGluAspTyrArgArgLysMetAlaGlySerAsnIleThrPhe 279

RESULT 25
US-10-199-329-8
; Sequence 8, Application US/10199329
; Patent No. 6953681
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-2
; CURRENT APPLICATION NUMBER: US/10/199,329
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27

```
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-329-8

Alignment Scores:
Pred. No.: 2,42e-32 Length: 283
Score: 419.50 Matches: 91
Percent Similarity: 53.0% Conservative: 57
Best Local Similarity: 32.6% Mismatches: 112
Query Match: 9.5% Indels: 19
DB: 2 Gaps: 4

US-10-768-158-1 (1-2419) x US-10-199-329-8 (1-283)
QY 81 GAGTTCATGCGGTGCGGCTCGCCCTTCTGCGCGGAAGATGGAGGATCGCCAAAC 140
DB 1 GluValAenGlyIleLeuMetSerLysLeuMetSerAspAsnTrpAspLysIleTrpAsn 20
QY 141 TTCCCGGTGCGCGGACGAGTGGGATCGTCACTACCCCAAGTCCCGGACCGAGCTTG 200
DB 21 PheGlnAlaLysProAspAspLeuLeuAlaThrTyrAlaLysAlaGlyThrTrp 40
QY 201 CTGCAGGAGGTGTCTACTTGTGTGAGCCAGGCGCTGACCCGATGAGATCGGCTTGATG 260
DB 41 ThrGlnIleValAspMetIleGlnAsnAspGlyAspValGlnLysCysGlnArgAla 60
QY 261 AACATCGACGACGAGTCCCGGTCTCGAGTAC-----CCACAGCGC-----GGC 305
DB 61 AsnThrTyrAspArgHisProPheIleGluTrpThrLeuProSerProLeuAsnSerGly 80
QY 306 CTGCACATCATCAGGAACGACCTCTCCCGCTCATCAAGGACCTCCCTACCGC 365
DB 81 LeuAspLeuAlaAsnLysMetProSerProThrLeuLysThrHisLeuProValHis 100
QY 366 TTTCGCGCTCTGACCTCCACATGGAGCTCCCAAGGTCACTATATGGCTCGCAACCCC 425
DB 101 MetLeuProProSerPheTrpLysGluAsnSerLysIleIleTyrValAlaArgAla 120
QY 426 AAGGATCTGGTGTCTTATATCAGTTCCACGCTCTCTCGGACCATGAGTACCGA 485
DB 121 LysAspCysLeuValSerTyrTyrPheSerArgMetAsnLysMetLeuProAspPro 140

US-10-768-158-1 (1-2419) x US-10-199-329-9 (1-283)
QY 81 GAGTTCATGCGGTGCGGCTCGCCCTTCTGCGCGGAAGATGGAGGATCGCCAAAC 140
DB 1 GluValAenGlyIleLeuMetSerLysLeuMetSerAspAsnTrpAspLysIleTrpAsn 20
QY 141 TTCCCGGTGCGCGGACGAGTGGGATCGTCACTACCCCAAGTCCCGGACCGAGCTTG 200
DB 21 PheGlnAlaLysProAspAspLeuLeuAlaThrTyrAlaLysAlaGlyThrTrp 40
QY 201 CTGCAGGAGGTGTCTACTTGTGTGAGCCAGGCGCTGACCCGATGAGATCGGCTTGATG 260
DB 41 ThrGlnIleValAspMetIleGlnAsnAspGlyAspValGlnLysCysGlnArgAla 60
QY 261 AACATCGACGACGAGTCCCGGTCTCGAGTAC-----CCACAGCGC-----GGC 305
DB 61 AsnThrTyrAspArgHisProPheIleGluTrpThrLeuProSerProLeuAsnSerGly 80
QY 306 CTGCACATCATCAGGAACGACCTCTCCCGCTCATCAAGGACCTCCCTACCGC 365
DB 81 LeuAspLeuAlaAsnLysMetProSerProThrLeuLysThrHisLeuProValHis 100
QY 366 TTTCGCGCTCTGACCTCCACATGGAGCTCCCAAGGTCACTATATGGCTCGCAACCCC 425
DB 101 MetLeuProProSerPheTrpLysGluAsnSerLysIleIleTyrValAlaArgAla 120
QY 426 AAGGATCTGGTGTCTTATATCAGTTCCACGCTCTCTCGGACCATGAGTACCGA 485
DB 121 LysAspCysLeuValSerTyrTyrPheSerArgMetAsnLysMetLeuProAspPro 140
```

```
QY 486 GGCACCTTTCAAGAAATTCGCGGAGGTTTATGAATGATAAGCTGGGCTACGGCTCTCTGG 545
Db 141 GlyThrLeuGlyGluTyrlleGluGlnPheLysAlaGlyLysValLeuTrpGlySerTrp 160
QY 546 TTTGAGCAGTCGACGAGTTCCTGGAGCAGCCGATGGACTCGAAGCTGCTTTTCTCAAG 605
Db 161 TyrAspHisValLysGlyTrpTrpAspValLysAspGlnHisArgIleLeuTyrllePhe 180
QY 606 TATGAAGACATGCATCGGACCTGCTGACGATGGTGAGCAGCTGGCCAGATTCCTGGGG 665
Db 181 TyrGluAspMetLysGluAspProLysArgGluIleLysLysIleAlaLysPheLeuGlu 200
QY 666 GTGTCCTGTGACAGGCCAGCTGGAAGCCCTGACGAGCAGCTGCCACCACTGGTG--- 722
Db 201 LysAspIleSerGluGluValLeuAsnLysIleIleTyrlleHisThrSerPheAspValMet 220
QY 723 -----GACCAGTGTGCAAGCTGAGCGCTGAGCCCTGCCC----- 752
Db 221 LysGluAsnProMetAlaAsnTyrlleThrThrLeuProSerSerIleMetAspHisSerIle 240
QY 753 -----GTGGCCGGGGAAGAGTTCGGCTGTGGAAGGACATCTTACCGTCTCCATG 803
Db 241 SerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrllePheThrValAlaGln 260
QY 804 AATGAGAAGTTTGACTTGGTGATATAACAGAGATGGGAAGTGTGACCTCACGTTT 860
Db 261 SerGluAspPheAspGluAspTyrlleArgArgLysMetAlaGlySerAsnIleThrPhe 279

RESULT 27
US-10-199-330-8
; Sequence 8, Application US/10199330
; Patent No. 6967094
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-1
; CURRENT APPLICATION NUMBER: US/10/199,330
; CURRENT FILING DATE: 2002-07-22
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-330-8

Alignment Scores:
Pred. No.: 2,42e-32 Length: 283
Score: 419.50 Matches: 91
Percent Similarity: 53.0% Conservative: 57
Best Local Similarity: 32.6% Mismatches: 112
Query Match: 9.5% Indels: 19
DB: 2 Gaps: 4

US-10-768-158-1 (1-241) x US-10-199-330-8 (1-283)
QY 81 GAGTTCCATGGCGTGGCTGCGCCCTTCTGCGCGGGAAGATGAGGAGATGCCCAAC 140
Db 1 GluValAsnGlyIleLeuMetSerLysLysLeuMetSerAspAsnTrpAspLysIleTrpAsn 20
QY 141 TTCGCGTGGCGGCCAGGACCTGTGATCGTACCTACCCCAAGTCGGGACCACTG 200
Db 21 PheGlnAlaLysProAspAspLeuIleAlaThrTyrlleAlaLysAlaGlyThrThrTrp 40
```

```
QY 201 CTGACGAGAGTGTCTACTTGGTGAGCGAGGCGCTGACCCCGATGAGATCGGCTTGATG 260
Db 41 ThrGlnGluIleValAspMetIleGlnAsnAspGlyAspValGlnLysCysGlnArgAla 60
QY 261 AACATCGACGAGCAGCTCCGGTCTCTGAGTAC-----CCACAGCCG-----GGC 305
Db 61 AsnThrTyrAspArgHisProPheIleGluTrpThrLeuProSerProLeuAsnSerGly 80
QY 306 CTGACATCATCAAGAACTGACCTCTCCCGCCTCATCAAGAGCCACCTGCCCTACCGC 365
Db 81 LeuAspLeuAlaAsnLysMetProSerProArgThrLeuLysThrHisLeuProValHis 100
QY 366 TTTTGGCCTCTGACCTCCACATGAGACTCCAAAGTTCATCTATATGCTGCGCAACCCC 425
Db 101 MetLeuProProSerPheTrpLysGluAsnSerLysIleIleTyrlleValAlaArgAsnAla 120
QY 426 AAGCATCTGTGGTGTCTTATTATCAGTTCACCGCTCTCTGCGGACCATGAGCTACCGA 485
Db 121 LysAspCysLeuValSerTyrllePheSerArgMetAsnLysMetLeuProAspPro 140
QY 486 GGCACCTTTTCAAGAAATTCGCGGAGGTTTATGAATGATAAGCTGGGCTCGCTCTGG 545
Db 141 GlyThrLeuGlyGluTyrlleGluGlnPheLysAlaGlyLysValLeuTrpGlySerTrp 160
QY 546 TTTGAGCAGTCGACGAGTTCCTGGAGCAGCCGATGGACTCGAAGCTGCTTTTCTCAAG 605
Db 161 TyrAspHisValLysGlyTrpTrpAspValLysAspGlnHisArgIleLeuTyrllePhe 180
QY 606 TATGAAGACATGCATCGGACCTGCTGACGATGGTGAGCAGCTGGCCAGATTCCTGGGG 665
Db 181 TyrGluAspMetLysGluAspProLysArgGluIleLysLysIleAlaLysPheLeuGlu 200
QY 666 GTGTCCTGTGACAGGCCAGCTGGAAGCCCTGACGAGCAGCTGCCACCACTGGTG--- 722
Db 201 LysAspIleSerGluGluValLeuAsnLysIleIleTyrlleHisThrSerPheAspValMet 220
QY 723 -----GACCAGTGTGCAAGCTGAGCGCTGAGCCCTGCCC----- 752
Db 221 LysGluAsnProMetAlaAsnTyrlleThrThrLeuProSerSerIleMetAspHisSerIle 240
QY 753 -----GTGGCCGGGGAAGAGTTCGGCTGTGGAAGGACATCTTACCGTCTCCATG 803
Db 241 SerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrllePheThrValAlaGln 260
QY 804 AATGAGAAGTTTGACTTGGTGATATAACAGAGATGGGAAGTGTGACCTCACGTTT 860
Db 261 SerGluAspPheAspGluAspTyrlleArgArgLysMetAlaGlySerAsnIleThrPhe 279
```

RESULT 28

```
US-10-199-330-9
; Sequence 9, Application US/10199330
; Patent No. 6967094
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-1
; CURRENT APPLICATION NUMBER: US/10/199,330
; CURRENT FILING DATE: 2002-07-22
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 283
```


Db 221 LysGlnAsnProMetAlaAsnTyrThrThrLeuProSerSerIleMetAspHisSerIle 240
QY 753 -----GTGGCGCGGGAAGAGTGGGCTGTGGAAGGACATCTTCACCGTCTCCATG 803
Db 241 SerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAlaGln 260
QY 804 AATGAGAAGTTTGACTTGGTGTATTAACAGAGATGGGAAAGTGTGAGCCTCACGTTT 860
Db 261 SerGluAspPheAspGluAspTyrArgLysLysMetAlaGlySerThrIleThrPhe 279
RESULT 33
US-09-609-816-6
; Sequence 6, Application US/09609816
; Patent No. 6436684
; GENERAL INFORMATION:
; APPLICANT: Woodage, Trevor
; APPLICANT: Wei, Minh Hui
; APPLICANT: Kodira, Chinappa
; APPLICANT: Beasley, Ellen
; APPLICANT: DiFrancesco, Valentina
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669PCT
; CURRENT APPLICATION NUMBER: US/09/609,816
; CURRENT FILING DATE: 2000-07-03
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 304
; TYPE: PRT
; ORGANISM: HUMAN
US-09-609-816-6
Alignment Scores:
Pred. No.: 4.33e-31 Length: 304
Score: 407.00 Matches: 94
Percent Similarity: 51.7% Conservative: 61
Best Local Similarity: 31.3% Mismatches: 125
Query Match: 9.2% Indels: 20
DB: 2 Gaps: 5
US-10-768-158-1 (1-2419) x US-09-609-816-6 (1-304)
QY 21 ATGGCGGAGACGAGCGCGGAGACCCAGACCCCGGGGGAGTTCGAG--AGCAAGTAC 77
Db 1 MetAlaLysIleGluLysAsnAlaProThrMetGluLysLysProGluLeuPheAsnIle 20
QY 78 TTCGAGTTCCATGGCGTGGCGTGGCGCCCTTCTGCCCGCGGGAAGATGGAGGAGATGCC 137
Db 21 MetGluValAspGlyValProThrLeuIleLeuSerLysGluTyrTrpGluLysValCys 40
QY 138 AACTTCCCGTGGCGGCCAGGACGAGTGTGGATCGTCACCTACCCCAAGTCGGGACGACG 197
Db 41 AsnPheGlnAlaLysProAspAspLeuIleLeuAlaThrTyrProLysSerGlyThrThr 60
QY 198 TTGCTGCAGGAGGTGGTCTACTTGTGTAGCCAGGCGGTGACCCCGATGAGATCGGCTTG 257
Db 61 TrpMetHisGluIleLeuAspMetIleLeuAsnAspGlyAspValGluLysCysLysArg 80
QY 258 ATGAACATCGACGACGACCTCCGCTCTGGAG-----TACCCA-----CAGCCG 302
Db 81 AlaGlnThrLeuAspArgHisAlaPheLeuGluLeuLysPheProHisLysGluLysPro 100
QY 303 GGCTTGGACATCATCAAGAACTGACCTCTCCCGCTCATCAAGAGCACCTGCCCTAC 362
Db 101 AspLeuGluPheValLeuMetSerSerProGlnLeuIleLysThrHisLeuProSer 120

QY 363 CGCTTTCTGCTCTGACCTCCCAATGGAGACTCCCAAGGTGCATCTATATGGCTCGCAAC 422
Db 121 HisLeuIleProProSerIleTrpLysGluAsnCysLysIleValTyrValAlaArgAsn 140
QY 423 CCAAGATCTGGTGGTGTCTTATATCAGTTCCACCGCTCTCTGCGGACCATGAGCTAC 482
Db 141 ProllysAspCysLeuValSerTyrHisPheHisArgMetAlaSerPheMetProAsp 160
QY 483 CGAGGCACCTTCAAGAAATCTCGCGGAGGTTTATGAATGATAAGCTGGGCTACGGCTCC 542
Db 161 ProGlnAsnLeuGluGluPheTyrGluLysPheMetSerGlyLysValValGlyGlySer 180
QY 543 TGGTTTGAGCAGCTGCAGGAGTTCTGGGAGCACCCTGAGCTCGAAGCTGCTTTTCTC 602
Db 181 TrpPheAspHisValLysGlyTyrTrpAlaAlaLysAspThrHisArgIleLeuTyrLeu 200
QY 603 AAGTATGAAGACATGCTCGGACCTGGTGTGAGTGGTGGAGCAGCTGCCAGATCTCTG 662
Db 201 PheTyrGluAspIleLysLysAsnProLysHisGluIleHisLysValLeuGluPheLeu 220
QY 663 GGGGTGCTGTGACAAAGGCCACCTGGAAGCCCTGACGGAGCACTGCCACCAGCTGTG 722
Db 221 GluLysThrLeuSerGlyAspValIleAsnLysIleValHisHisThrSerPheAspVal 240
QY 723 -----GACCAGTGTCTGCACGCTGAGGCCCTGCCCC----- 752
Db 241 MetLysAspAsnProMetAlaAsnHisThrAlaValProAlaHisIlePheAsnHisSer 260
QY 753 -----GTGGCGCGGGAAGAGTGGGCTGTGGAAGGACATCTTCACCGTCTCC 800
Db 261 IleSerLysPheMetArgLysGlyMetProGlyAspTrpLysAsnHisPheThrValAla 280
QY 801 ATGAATGAGAAGTTTGACTTGGTGTATATAACAGAAAGTGGAAAGTGTGACCTCACGTTT 860
Db 281 MetAsnGluAsnPheAspLysHisTyrGluLysLysMetAlaGlySerThrLeuAsnPhe 300
RESULT 34
US-10-199-334-6
; Sequence 6, Application US/10199334
; Patent No. 6905855
; GENERAL INFORMATION:
; APPLICANT: Woodage, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669Div-3
; CURRENT APPLICATION NUMBER: US/10/199,334
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-334-6
Alignment Scores:
Pred. No.: 4.33e-31 Length: 304
Score: 407.00 Matches: 94
Percent Similarity: 51.7% Conservative: 61
Best Local Similarity: 31.3% Mismatches: 125
Query Match: 9.2% Indels: 20
DB: 2 Gaps: 5

Db 221 GlulysThrLeuSerGlyAspValIleAsnLysIleValHisHisThrSerPheAspVal 240
Qy 723 -----GACCAGTGTGCAACGCTGAGCGCTGCGCC----- 752
Db 241 MetLysAspAsnProMetAlaAsnHisThrAlaValProAlaHisIlePheAsnHisSer 260
Qy 753 -----GTGGCGGGGAGAGATTGGGCTGTGGAAGGACATCTTCACCGTCTCC 800
Db 261 IleSerLysPheMetArgLysGlyMetProGlyAspTrpLysAsnHisPheThrValAla 280
Qy 801 ATGAATGAGAAGTTGACTTGGTGATATAACAGAAAGATGGGAAAGTGTCACCTCAGCTTT 860
Db 281 MetAsnGluAsnPheAspLysHisTyrGluLysLysMetAlaGlySerThrLeuAsnPhe 300
RESULT 36
US-10-199-330-6
; Sequence 6, Application US/10199330
; Patent No. 6967094
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-1
; CURRENT APPLICATION NUMBER: US/10/199,330
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-330-6
Alignment Scores:
Pred. No.: 4.33e-31 Length: 304
Score: 407.00 Matches: 94
Percent Similarity: 51.7% Conservative: 61
Best Local Similarity: 31.3% Mismatches: 125
Query Match: 9.2% Indels: 20
DB: 2 Gaps: 5
US-10-768-158-1 (1-2419) x US-10-199-330-6 (1-304)
Qy 21 ATGGCGGAGACGAGCGCGAGACCCCGAGCCCGGGGAGTTCGAG---AGCAAGTAC 77
Db 1 MetAlaLysIleGluLysAsnAlaProThrMetGluLysLysProGluLeuPheAsnIle 20
Qy 78 TTCGAGTTCCATGGCGTGGCGCTGCGCGCTTCTGCGCGGGAAGATGGAGGAGATCGCC 137
Db 21 MetGluValAspGlyValProThrLeuIleLeuSerLysGluTrpTrpGluLysValCys 40
Qy 138 AACTTCCGGTGGCGGCCAGGACGTGTGGATGTGTCACCTACCCCAAGTCCGGCACACAGC 197
Db 41 AsnPheGlnAlaLysProAspAspIleLeuAlaThrPrProLysSerGlyThrThr 60
Qy 198 TTGCTGCGAGAGGTGGTCTACTTGGTGAGCCAGGCGCTGACCCCGATGAGATCGGCTTG 257
Db 61 TrpMetHisGluIleLeuAspMetIleLeuAsnAspGlyAspValGluLysCysLysArg 80
Qy 258 ATGAACATCGACGAGCTCCCGTCTCGGAG-----TACCCA-----CAGCCG 302
Db 81 AlaGlnThrLeuAspArgHisAlaPheLeuGluLeuLysPheProHisLysGluLysPro 100
Qy 303 GGCCTGGACATCATCAAGAACTGACCTCTCCCGCTCATCAAGAGCCACCTGCCCTTAC 362

Db 101 AspLeuGluPheValLeuMetSerSerProGlnLeuIleLysThrHisLeuProSer 120
Qy 363 CGCTTTCTGCGCTCTGACCTCCACAATGGAGATCCCAAGGTCATCTATATGGCTCGCAAC 422
Db 121 HisLeuIleProProSerIleTrpLysGluAsnCysLysIleValTyrValAlaArgAsn 140
Qy 423 CCAAGGATCTGGTGGTGTCTTATATCAGTTCCACCGCTCTCTGCGGACCATCAGGTAC 482
Db 141 ProLysAspCysLeuValSerTyrHisPheHisArgMetAlaSerPheMetProAsp 160
Qy 483 CGAGGCACCTTTCAGAAATCTGCGGAGGTTTATGAATGATAAGCTGGGCTACGGCTCC 542
Db 161 ProGlnAsnLeuGluGluPheTyrGluLysPheMetSerGlyLysValValGlyGlySer 180
Qy 543 TGGTTTGAGCACGTCGAGGAGTTCTGGGAGCACCGCATGGACTCGAAGCTCTTTTCTC 602
Db 181 TrpPheAspHisValLysGlyTrpTrpAlaAlaLysAspThrHisArgIleLeuTyrLeu 200
Qy 603 AAGTATGAAGACATGTCATCGGACCTGGTGACGATGGTGAGCAGCTGGCCAGATTCTGT 662
Db 201 PheTyrGluAspIleLysLysAsnProLysHisGluIleHisLysValLeuGluPheLeu 220
Qy 663 GGGGTGCTCTGTGACAAGGCCACGCTGGAAGCCCTGACGGAGCACTGCCACCAGCTGTG 722
Db 221 GluLysThrLeuSerGlyAspValIleAsnLysIleValHisHisThrSerPheAspVal 240
Qy 723 -----GACCAGTGTGCAACGCTGAGCGCTGCGCC----- 752
Db 241 MetLysAspAsnProMetAlaAsnHisThrAlaValProAlaHisIlePheAsnHisSer 260
Qy 753 -----GTGGCGCGGGGAGAGTGGGCTGTGGAAGGACATCTTCACCGTCTCC 800
Db 261 IleSerLysPheMetArgLysGlyMetProGlyAspTrpLysAsnHisPheThrValAla 280
Qy 801 ATGAATGAGAAGTTGACTTGGTGATATAACAGAAAGATGGGAAAGTGTCACCTCAGCTTT 860
Db 281 MetAsnGluAsnPheAspLysHisTyrGluLysLysMetAlaGlySerThrLeuAsnPhe 300
RESULT 37
US-09-609-816-4
; Sequence 4, Application US/09609816
; Patent No. 6436684
; GENERAL INFORMATION:
; APPLICANT: Woodage, Trevor
; APPLICANT: Wei, Minh Hui
; APPLICANT: Kodira, Chinappa
; APPLICANT: Beasley, Ellen
; APPLICANT: DiFrancesco, Valentina
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669PCT
; CURRENT APPLICATION NUMBER: US/09/609,816
; CURRENT FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 304
; TYPE: PRT
; ORGANISM: HUMAN
US-09-609-816-4
Alignment Scores:
Pred. No.: 1.7e-30 Length: 304
Score: 401.00 Matches: 93
Percent Similarity: 51.7% Conservative: 62

```
Best Local Similarity: 31.0% Mismatches: 125
Query Match: 9.1% Indels: 20
DB: 2 Gaps: 5

US-10-768-158-1 (1-2419) x US-09-609-816-4 (1-304)

QY 21 ATGCGGAGAGCGAGCGGAGACCCAGCAGCCCGGGGGAGTTCGAG---AGCAAGTAC 77
DB 1 MetAlaLysIleGluLysAsnAlaProThrMetGluLysLysProGluLeuPheAsnIle 20

QY 78 TTCGAGTTCCATGCGTGGCGTCCGCCCTTCTGCGCGGGAAGATGGAGGATCGCC 137
DB 21 MetGluValAspGlyValProThrLeuLeuSerLysGluTrpTpgLysValCys 40

QY 138 AACTTCCGGTGGCGGCGGAGCGTGTGGATCGTACCTACCCCAAGTCCGGCAGCAGC 197
DB 41 AsnPheGlnAlaLysProAspAspLeuLeuAlaThrTyrProLysSerGlyThr 60

QY 198 TTGCTGACGAGGTTGCTACTTGTGAGCGCGCTGACCCCGATGAGATCGGCTTG 257
DB 61 TrpMetHisGluIleLeuAspMetIleuAsnAspGlyAspValGluLysCysLysArg 80

QY 258 ATGAACATCGACGAGCGTCCCGTCTCGAG-----TACCCA-----CAGCCG 302
DB 81 AlaGlnThrLeuAspArgHisAlaPheLeuGluLysPheProHisLysGluLysPro 100

QY 303 GGCCTGGACATCATCAAGGAAGTACCTCTCCCGCTCATCAAGAGCCACTGCCCTAC 362
DB 101 AspLeuGluPheValLeuMetSerProGlnLeuIleLysThrHisLeuProSer 120

QY 363 CGCTTTCTGCGCTCTGACCTCCACAATGGAGACTCCCAAGGTCTATATATGCTCGCAAC 422
DB 121 HisLeuIleProProSerIleTrpLysGluAsnCysLysIleValTyrValAlaArgAsn 140

QY 423 CCCAAGGATCTGGTGGTGTCTTATTATCAGTTCCACCGCTCTCTCGCGACCATGAGTAC 482
DB 141 ProLysAspCysLeuValSerTyrTyrHisPheHisArgMetAlaSerPheMetProAsp 160

QY 483 CGAGGCACCTTCAAGATTTCTCGCGAGGTTTATGATGATAGCTGGGTACGGCTCC 542
DB 161 ProGlnAsnLeuGluPheTyrGluLysPheMetSerGlyLysValValGlySer 180

QY 543 TGGTTTGGACGCTGAGGAGTCTTGGAGCAGCGCATGGACTCGAAGCTGCTTTTCTC 602
DB 181 TrpPheAspHisValLysGlyTrpTrpAlaLysAspMetHisArgIleLeuTyrLeu 200

QY 603 AAGTATGAAGACATGCAATGCGGAGCTGTCAGTGTGGAGCAGCTGGCGAGATTCTCTG 662
DB 201 PheTyrGluAspIleLysLysAsnProLysHisGluIleHisLysValLeuGluPheLeu 220

QY 663 GGGGTGCTCTGTGACAGCCGCTGGAAGCCCTGACGCGGAGCAGTGCACCACTGGTG 722
DB 221 GluLysThrTrpSerGlyAspValIleAsnLysIleValHisHisThrSerPheAspVal 240

QY 723 -----GACCAGTGTGCAACCTGAGCGCCCTGCGCC-----752
DB 241 MetLysAspAsnProMetAlaAsnHisThrAlaValProAlaHisIlePheAsnHisSer 260

QY 753 -----GTGGCGCGGAGAGTGTGGCTGTGGAAGGACATCTTCAACGCTTCC 800
DB 261 IleSerLysPheMetArgLysGlyMetProGlyAspTrpLysAsnHisPheThrValAla 280

QY 801 ATGATGAGAGTGTGCTGTGATAAACAGAGAGATGGGAAAGTGAGCTCACGTTT 860
DB 281 LeuAsnGluAsnPheAspLysHisTyrGluLysLysMetAlaGlySerThrLeuAsnPhe 300

RESULT 38
US-10-199-334-4
; Sequence 4, Application US/10199334
; Patent No. 6905855
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
```

```
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-3
; CURRENT APPLICATION NUMBER: US/10/199,334
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-199-334-4
```

```
Alignment Scores:
Pred. No.: 1.7e-30 Length: 304
Score: 401.00 Matches: 93
Percent Similarity: 51.7% Conservative: 62
Best Local Similarity: 31.0% Mismatches: 125
Query Match: 9.1% Indels: 20
DB: 2 Gaps: 5
```

US-10-768-158-1 (1-2419) x US-10-199-334-4 (1-304)

```
Db 201 PheTyrGluAspIleLysLysAsnProLysHisGluIleHisLysValLeuGluPheLeu 220
Qy 663 GGGGTGCTCTGTGACAAAGCCAGCTGGAAGCCCTGACGGAGCACTGCCACCAAGCTGGTG 722
Db 221 GluLysThrTrpSerGlyAspValIleAsnLysIleValHisHisThrSerPheAspVal 240
Qy 723 -----GACCAGTGTGCACGCTGAGCCCTGCCC----- 752
Db 241 MetLysAspAsnProMetAlaAsnHisThrAlaValProAlaHisIlePheAsnHisSer 260
Qy 753 -----GTGGCGCGGGAGAGCTTGGGCTGTGGAAGGACATCTTCACCGTCTCC 800
Db 261 IleSerLysPheMetArgLysGlyMetProGlyAspTrpLysAsnHisPheThrValAla 280
Qy 801 ATGAATGAGAAGTTTGACTTGGTGATAAACAGAGATGGAAAGTGATGACCTCACGTTT 860
Db 281 LeuAsnGluAsnPheAspLysHisTyrGluLysLysMetAlaGlySerThrLeuAsnPhe 300
RESULT 39
US-10-199-329-4
; Sequence 4, Application US/10199329
; Patent No. 6953681
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-2
; CURRENT APPLICATION NUMBER: US/10/199,329
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-329-4
Alignment Scores:
Pred. No.: 1.7e-30 Length: 304
Score: 401.00 Matches: 93
Percent Similarity: 51.7% Conservative: 62
Best Local Similarity: 31.0% Mismatches: 125
Query Match: 9.1% Indels: 20
DB: 2 Gaps: 5
US-10-768-158-1 (1-2419) x US-10-199-329-4 (1-304)
Qy 21 ATGGCGGAGCGAGCGCGAGACCCCGAGCCCGGGGGAGTTCGAG---AGCAAGTAC 77
Db 1 MetAlaLysIleGluLysAsnAlaProThrMetGluLysLysProGluLeuPheAsnIle 20
Qy 78 TTCGAGTTCATGGGTGGCGGCTGCGGCCCTTTCGCGCGGAGAGATGGAGGATCGCC 137
Db 21 MetGluValAspGlyValProThrLeuIleLeuSerLysGluIleTrpTrpGluLysValCys 40
Qy 138 AACTTCCCGGTGCGCGCGAGCGAGCTGGGATCGTCACCTACCCCAAGTCGCGGACCGAGC 197
Db 41 AsnPheGlnAlaLysProAspAspLeuLeuAlaThrTyrProLysSerGlyThrThr 60
Qy 198 TTGCTGAGGAGGTGGTCTACTTGTGTGAGCCAGGCGCGCTGACCCCGATGAGATCGGCTTG 257
Db 61 TrpMetHisGluIleLeuAspMetIleLeuAsnAspGlyAspValGluLysCysLysArg 80
Qy 258 ATGAACATCGACGACAGCTCCCGTCTCGAG-----TACCACA-----CAGCCG 302
```

```
Db 81 AlaGlnThrLeuAspArgHisAlaPheLeuGluLeuLysPheProHisLysGluLysPro 100
Qy 303 GGCTTGACATCATCAAGGAACACTGACCTCTCCCGCCTCATCAAGAGCCACTGCCCTTAC 362
Db 101 AspLeuGluPheValLeuGluMetSerProGlnLeuIleLysThrHisLeuProSer 120
Qy 363 CGCTTTTCGCCTCTGACCTCCACAATGGAGACTCCCAAGGTCACTATATATGGCTCGCAAC 422
Db 121 HisLeuIleProProSerIleTrpLysGluAsnCysLysIleValTyrValAlaArgAsn 140
Qy 423 CCCAAGGATCTGGTGGTGTCTTATATACAGTTCACCGCTCTCTGCGGACCATGAGTAC 482
Db 141 ProLysAspCysLeuValSerTyrTyrHisPheHisArgMetAlaSerPheMetProAsp 160
Qy 483 CGAGGCACTTTCAGAAATCTCTGCCGAGGTTTATGAATGATAGCTGGGTACGGCTCC 542
Db 161 ProGlnAsnLeuGluPheTyrGluLysPheMetSerGlyLysValValGlyGlySer 180
Qy 543 TGGTTTGAGCAGCTGCGAGGAGTTCTGGAGACCCGCGATGGACTCGAAGTGTCTTTTCTC 602
Db 181 TrpPheAspHisValLysGlyTrpTrpAlaAlaLysAspMetHisArgIleLeuTyrLeu 200
Qy 603 AAGTATGAACATGTCATCGGACCTGGTGACGATGTGGAGCAGCTGCCAGATTCTGTG 662
Db 201 PheTyrGluAspIleLysLysAsnProLysHisGluIleHisLysValLeuGluPheLeu 220
Qy 663 GGGGTGTCTGTGTACAAGGCCAGCTGGAAGCCCTGACGGAGCACTGCCACCACTGGTG 722
Db 221 GluLysThrTrpSerGlyAspValIleAsnLysIleValHisHisThrSerPheAspVal 240
Qy 723 -----GACCAGTCTGCAACGCTGAGCGCTGCC----- 752
Db 241 MetLysAspAsnProMetAlaAsnHisThrAlaValProAlaHisIlePheAsnHisSer 260
Qy 753 -----GTGGCGCGGGAGAGTGGGTGTGGAGGACATCTTCACCGTCTCC 800
Db 261 IleSerLysPheMetArgLysGlyMetProGlyAspTrpLysAsnHisPheThrValAla 280
Qy 801 ATGAATGAGAAGTTTGACTTGGTGTATATAACAGAAAGATGGAAAGTGTGACCTCACGTTT 860
Db 281 LeuAsnGluAsnPheAspLysHisTyrGluLysLysMetAlaGlySerThrLeuAsnPhe 300
RESULT 40
US-10-199-330-4
; Sequence 4, Application US/10199330
; Patent No. 6967094
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-1
; CURRENT APPLICATION NUMBER: US/10/199,330
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-330-4
Alignment Scores:
Pred. No.: 1.7e-30 Length: 304
```



```
Score: 401.00 Matches: 93
Percent Similarity: 51.7% Conservative: 62
Best Local Similarity: 31.0% Mismatches: 125
Query Match: 9.1% Indels: 20
DB: 2 Gaps: 5

US-10-768-158-1 (1-2419) x US-10-199-330-4 (1-304)
Qy 21 ATGGCGAGAGCGGCGGAGACCCCGAGACCCCGGGGGAGTTCGAG---AGCAAGTAC 77
Db 1 MetAlaIysAlaIysAsnAlaProThrMetGluLysLysProGluLeuPheAsnIle 20
Qy 78 TTCAGATTCCATGCGTGGCGGCTCGCGCCCTTCTGCGCGGGAAGATGGAGGATCGCC 137
Db 21 MetGluValAspGlyValProThrLeuLeuLeuSerLysGluTrpTrpGluLysValCys 40
Qy 138 AACTTCCCGTGGCGGCGGAGAGTGTGGATCTGCATCCCAAGTCCCGGACCCAGC 197
Db 41 AsnPheGlnAlaLysProAspAspLeuLeuAlaThrTyrProLysSerGlyThrThr 60
Qy 198 TTGCTGAGGAGGTGCTACTTGTGTAGCGAGCGGCTGACCCCGATGATGCGGCTTG 257
Db 61 TrpMetHisGluLeuLeuAspMetIleLeuAsnAspGlyAspValGluLysCysLysArg 80
Qy 258 ATGAACATCGACGACGCTCCGCTCTGCGAG-----TACCCA-----CAGCG 302
Db 81 AlaGlnThrLeuAspArgHisAlaPheLeuLeuLysPheProHisLysGluLysPro 100
Qy 303 GCGCTGGACATCATCAAGAACTACCTCTCCCGCTCATCAAGAGCACCTGCCCTAC 362
Db 101 AspLeuGluPheValLeuGluMetSerSerProGlnLeuIleLysThrHisLeuProSer 120
Qy 363 CGCTTTGCGCTCTGACCTCCAAATGGAGACTCCAAAGTCACTATATGGTCGCGAAC 422
Db 121 HisLeuIleProProSerIleTrpLysGluAsnCysLysIleValValAlaArgAsn 140
Qy 423 CCCAGGATCTGGTGTCTTATATCAGTTCCACGCTCTCTCGGACCATCAGCTAC 482
Db 141 ProLysAspCysLeuValSerTyrTrpHisPheHisArgMetAlaSerPheMetProAsp 160
Qy 483 CGAGGCACTTCAAGAACTTCCCGGAGGTTTATGAATGATGAAGCTGGGTACGGCTCC 542
Db 161 ProGlnAsnLeuGluGluPheTyrGluLysPheMetSerGlyLysValValGlyLys 180
Qy 543 TGGTTTGAAGACGTCAGGAGTCTCGGAGCAGCCGATGGACTGCAAGCTCTTTCTC 602
Db 181 TrpPheAspHisValLysGlyTrpTrpAlaLysAspMetHisArgIleLeuTyrLeu 200
Qy 603 AAGTATGAAGACATCATCGGACCTGCTGACGATGGTGGAGCAGCTGGCCAGATTCTG 662
Db 201 PheTyrGluAspIleLysLysAsnProLysHisGluIleHisLysValLeuGluPheLeu 220
Qy 663 GGGGTGCTCTGTACAAGGCCAGCTGGAAGCCCTGACGGAGCACTGCCACAGCTGGTG 722
Db 221 GluLysThrTrpSerGlyAspValIleAsnLysIleValHisHisThrSerPheAspVal 240
Qy 723 -----GACCACTGTGCAACCTGAGCGCCCTGCGCC----- 752
Db 241 MetLysAspAsnProMetAlaAsnHisThrAlaValProAlaHisIlePheAsnHisSer 260
Qy 753 -----GTGGCGCGGGAAGAGTGTGGCTGTGGAAGGACATCTTCACGCTGCC 800
Db 261 IleSerLysPheMetArgLysGlyMetProGlyLysPheAsnHisPheThrValAla 280
Qy 801 ATGAATGAGAAGTTTGACTTGTGTATAAACAGAGATGGGAAAGTGTGACCTCAGCTTT 860
Db 281 LeuAsnGluAsnPheAspLysHisTyrGluLysLysMetAlaGlySerThrLeuAsnPhe 300
```

RESULT 41

```
US-09-949-016-6526
; Sequence 6526, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
```

```
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CU001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6526
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6526
```

```
Alignment Scores:
Pred. No.: 1,74e-28 Length: 350
Score: 381.00 Matches: 94
Percent Similarity: 51.8% Conservative: 52
Best Local Similarity: 33.3% Mismatches: 108
Query Match: 8.7% Indels: 28
DB: 2 Gaps: 7
```

US-10-768-158-1 (1-2419) x US-09-949-016-6526 (1-350)

```
Qy 54 CCGGGGAGTTCGAGACCAAGTACTTCGAGTTCATCGCGTGGCGTGGCGCCCTTCTGC 113
Db 13 ProGlyGlu-----TyrPheArgTyrLysGlyVal-----ProPhePro 25
Qy 114 CGCGGG-----AAGATGGAGGAGATCGCC-----AACTTCCCGTGGCGGCC 155
Db 26 ValGlyLeuTyrSerLeuGluSerIleSerLeuAlaGluAsnThrGlnAspValArgAsp 45
Qy 156 AGCAAGCTGTGGATCGTCACCTACCCCAAGTCCGGCAGCAGTGTGTCGAGGAGTGGTC 215
Db 46 AspAspIlePheIleIleThrTyrProLysSerGlyThrThrTrpMetIleGluIle 65
Qy 216 TACTTGTGTGAGCCAGGCGCTGACCCCGATGAGATCGGCTTGATGAACATTCGACGAGCAG 275
Db 66 CysLeuIleLeuLysGluGlyAspProSerTrpIleArgSerValProIleTrpGluArg 85
Qy 276 CTCCCGTCTCTGGAGTACCACAGCCGGCTGGACATCATCAAGGAACCTGACCTCTCCC 335
Db 86 AlaProTrpCysGlu---ThrIleValGlyAlaPheSerLeuProAspGlnTyrSerPro 104
Qy 336 CGCTCATCAAGAGCCACCTGCCCTACCGCTTCTGCGCTCTGACCTCCACATGGAGAC 395
Db 105 ArgLeuMetSerSerHisLeuProIleGlnIlePheThrLysAlaPhePheSerSerLys 124
Qy 396 TCCAAGTGTCTATATGCTGCGCAACCCCAAGGATCTGGTGGTGTCTTATTATCAGTTC 455
Db 125 AlaLysValIleTyrMetGlyArgAsnProArgAspValValSerLeuTyrHisTyr 144
Qy 456 CACCGCTCTCTGCGGACCATGAGCTACCGAGGCACTTTTCAAGAAATTCGCGGAGGTTT 515
Db 145 SerLysIleAlaGlyGlnLeuLysAspProGlyThrProAspGlnPheLeuArgAspPhe 164
Qy 516 ATGAATATAAGCTGGCTACGCTCGCTGCTGGTTCGAGCAGCTGCAGGAGTTCGGGAGCAC 575
Db 165 LeuLysGlyGluValGlnPheGlySerTrpPheAspHisIleLysGlyTrpLeuArgMet 184
Qy 576 CGCATGCACTCGAAGCTGCTTTTCTCAAGTATGAAGACATGATCGGACCTGTGTGACG 635
Db 185 LysGlyLysAspAsnPheLeuPheIleThrTyrGluGluLeuGlnAspLeuGlnGly 204
Qy 636 ATGTGGAGCAGCTGGCCAGATTCCTGGGGGTGTCTGTGTGACAAGGCCACCTGGAAAGCC 695
Db 205 SerValGluArgIleCysGlyPheLeuGlyArgProLeuGlyLysGluAlaLeuGlySer 224
```



```
QY 696 CTGACGGAGCAC-----TGCCACCAGCTGGTGGACCAAGTGTGCTGCAACGCTGAGGCC 746
   :::
Db 225 ValValAlaHisSerThrPheSerAlaMetLysAlaAsnThrMetSerAsnTyrThrLeu 244
   :::
QY 747 CTGCCC-----GTGGCGCGGGAAGAGTTGGG 773
   :::
Db 245 LeuProProSerLeuLeuAspHisArgGlyAlaPheLeuArgLysGlyValCysGly 264
   :::
QY 774 CTGTGGGAAGGACATCTTCACCGTCTCCATGAATGAGAAAGTTTGACTTGGTGATAAAGAC 833
   :::
Db 265 AspTrpLysAsnHisPheThrValAlaGlnSerGluAlaPheAspArgAlaTyrArgLys 284
   :::
QY 834 AAGATG 839
   :::
Db 285 GlnMet 286

RESULT 42
US-09-949-016-11211
; Sequence 11211, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11211
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11211

Alignment Scores:
Pred. No.: 1,82e-28 Length: 377
Score: 381.00 Matches: 94
Percent Similarity: 51.8% Conservative: 52
Best Local Similarity: 33.3% Mismatches: 108
Query Match: 8.7% Indels: 28
DB: 2 Gaps: 7

US-10-768-158-1 (1-2419) x US-09-949-016-11211 (1-377)
QY 54 CCGGGGAGTTCGAGACGAGTACTTCAGTTCATGCGTGGCGTGGCCCTTCTGTC 113
   :::
Db 40 ProGlyGlu-----TyrPheArgTyrLysGlyVal-----ProPhePro 52
   :::
QY 114 CGCGGG-----AAGATGGAGGAGATCGCC-----AACTTCCCGTCCGCGCC 155
   :::
Db 53 ValGlyLeuTyrSerLeuGluSerIleSerLeuAlaGluAsnThrGlnAspValArgAsp 72
   :::
QY 156 AGCAGCTGTGGATCGTCACTACCCCAAGTCCGCGACCACTGCTGTCAGGAGGTGTC 215
   :::
Db 73 AspAspIlePheIleIleTyrTyrProLysSerGlyThrTrpMetIleGluIleIle 92
   :::
QY 216 TACTTGTGTGACCGGCGTGCACCCGATGAGATCGGCTTGATGAACATCGACGAGCAG 275
   :::
Db 93 CysLeuIleLeuLysGluGlyAspProSerTrpIleArgSerValProIleTrpGluArg 112
   :::
QY 276 CTCCCGGTCTGGAGTACCCACAGCCGCGCTGGACATCATCAAGGAAGTCACTCTCCC 335
   :::
Db 113 AlaProTrpCysGlu-----ThrIleValGlyAlaPheSerLeuProAspGlnTyrSerPro 131
   :::
QY 336 CGCCTCATCAAGAGCCACCTGCCCTTCTGACGCTTTCTGCCCTCTGACCTCCACAAATGGAGAC 395
```

```
Db 132 ArgLeuMetSerSerHisLeuProIleGlnIlePheThrLysAlaPhePheSerLys 151
   :::
QY 396 TCCAAGGTCACTATATGGCTCGCAACCCCAAGGATCTGGTGGTCTTATTATCAGTTC 455
   :::
Db 152 AlalysValIleTyrMetGlyArgAsnProArgPheValValSerLeuTyrHisTyr 171
   :::
QY 456 CACCGCTCTCTCGCGACCATGAGCTACCGAGGCACCTTTCAAGAATCTCGCGGAGTTT 515
   :::
Db 172 SerLysIleAlaGlyGlnLeuLysAspProGlyThrProAspGlnPheLeuArgAspPhe 191
   :::
QY 516 ATGAATGTAAGCTGGGCTACGGCTCTGGTGGTTCAGCAGCTGCAGGAGTTCTGGGAGCAC 575
   :::
Db 192 LeuLysGlyGluValGlnPheGlySerTrpPheAspHisIleLysGlyTrpLeuArgMet 211
   :::
QY 576 CGCATGGACTCGAAAGCTGCTTTTCTCAAGTATGAAGACATGCATCGGGACCTGTGACG 635
   :::
Db 212 LysGlyLysAspAsnPheLeuPheIleThrTyrGluGluGlnGlnAspLeuGlnGly 231
   :::
QY 636 ATGCTGGAGCAGCTGGCCAGATTCCTGGGGGTGCTCTGTGACAAGGCCAGCTGGAAGCC 695
   :::
Db 232 SerValGluArgIleCysGlyPheLeuGlyArgProLeuGlyLysGluAlaLeuGlySer 251
   :::
QY 696 CTGACGGAGCAC-----TGCACACAGCTGGTGGACCAAGTGTGCAACGCTGAGGCC 746
   :::
Db 252 ValValAlaHisSerThrPheSerAlaMetLysAlaAsnThrMetSerAsnTyrThrLeu 271
   :::
QY 747 CTGCCC-----GTGGCGCGGGAAGAGTTGGG 773
   :::
Db 272 LeuProProSerLeuLeuAspHisArgGlyAlaPheLeuArgLysGlyValCysGly 291
   :::
QY 774 CTGTGGGAAGGACATCTTCACCGTCTCCATGAATGAGAAAGTTTGACTTGGTGATAAAGAC 833
   :::
Db 292 AspTrpLysAsnHisPheThrValAlaGlnSerGluAlaPheAspArgAlaTyrArgLys 311
   :::
QY 834 AAGATG 839
   :::
Db 312 GlnMet 313

RESULT 43
US-09-949-016-8464
; Sequence 8464, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8464
; LENGTH: 264
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8464

Alignment Scores:
Pred. No.: 1,85e-28 Length: 264
Score: 380.00 Matches: 76
Percent Similarity: 57.7% Conservative: 48
Best Local Similarity: 35.3% Mismatches: 87
Query Match: 8.6% Indels: 4
DB: 2 Gaps: 2

US-10-768-158-1 (1-2419) x US-09-949-016-8464 (1-264)
```

```
QY 72 AAGTACTTCGAGTTCATGGCGTGGCGTCCGCCCTTCTGCGCGGGAAGATGGAGGAG 131
Db 14 LysLeuLysGluValGluGlyThrLeuLeuGlnProAlaThrValAspAsnTrpSerGln 33
QY 132 ATCCCAACTTCCCGGTGCGGCCCGGACGATGCGTGGATCGTCACCTACCCCAAGTCGCGC 191
Db 34 IleglnSerPheGluAlaLysProAspAspLeuLeuLeuCysThrTyProLysAlaGly 53
QY 192 ACCAGCTTGTGCGAGGAGTGTCTACTTGTGTAGCAGCGCGCTGACCCCGATGAGATC 251
Db 54 ThrThrTrpIleGlnGluValAspMetIleGlnAsnGlyAspValGluLysCys 73
QY 252 GGCTTGATGACATCGAGCAGCAGCTCCGCTCTGGAGTAC-----CCACAGCG 302
Db 74 GlnArgAlaIleIleGlnHisArgHisPheIleGluTrpAlaArgProProGlnPro 93
QY 303 ---GGCTGGACATCATCAAGAACTGACCTCTCCCGCTCATCAAGAGCCACTGCCC 359
Db 94 SerGlyValGluLysAlaLysAlaMetProSerProArgIleLeuLysThrHisLeuSer 113
QY 360 TACCGTTCCTGCGCTGACCTCCACATGGAGACTCCCAAGTCATCTATATGCTGCG 419
Db 114 ThrGlnLeuLeuProProSerPheTrpGluAsnAsnCysLysPheLeuTyValAlaArg 133
QY 420 AACCCCAAGATCGTGGTGTCTTATATCAGTTCACCGCTCTCTCGGACCATGAGC 479
Db 134 AsnAlaLysAspCysMetValSerTyThrHisPheGlnArgMetAsnHisMetLeuPro 153
QY 480 TACCGAGCAGCTTCAAGAACTTCGCGGAGTTTATGATGATGATAAGCTGGGTACGGC 539
Db 154 AspProGlyThrTrpGluTyPheGluThrPheIleAsnGlyLysValValTrpGly 173
QY 540 TCCTGGTTGAGCAGCTGCGAGGTTCTGGAGCAGCCGCTGAGCTCGAAGCTGCTTTT 599
Db 174 SerTrpPheAspHisValLysGlyTrpTrpGluMetLysAspArgHisGlnIleLeuPhe 193
QY 600 CTCAAGTATGAAGATCATCGGAGCTGAGCTGAGCTGGAGCAGCTGGCAGATTC 659
Db 194 LeuPheTyGluAspIleLysArgAspProLysHisGluIleArgLysValMetGlnPhe 213
QY 660 CTGGGGGTCTCTGTGACAGGCCAGCTGGAAGCCCTGACGGAG 704
Db 214 MetGlyLysLysValAspGluThrValLeuAspLysIleValGln 228
```

RESULT 44

```
US-08-852-481-2
; Sequence 2, Application US/08852481
; Patent No. 5928931
; GENERAL INFORMATION:
; APPLICANT: Grun Ph.D., Felix
; APPLICANT: Buck Ph.D., Jochen
; TITLE OF INVENTION: ISOLATION, PURIFICATION AND CLONING OF
; TITLE OF INVENTION: RETINOL DEHYDRATASE
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/852.481
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
```

```
; APPLICATION NUMBER: US 60/017,178
; FILING DATE: 09-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldman, Michael L.
; REGISTRATION NUMBER: 30,727
; REFERENCE/DOCKET NUMBER: 19603/771
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 263-1304
; TELEFAX: (716) 263-1600
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 350 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-852-481-2
Alignment Scores:
Pred. No.: 9,56e-24 Length: 350
Score: 333.00 Matches: 88
Percent Similarity: 46.3% Conservative: 64
Best Local Similarity: 26.8% Mismatches: 106
Query Match: 7.6% Indels: 70
DB: Gaps: 10
US-10-768-158-1 (1-2419) x US-08-852-481-2 (1-350)
QY 57 GGGGAGTTCGAGAGCAAGTACTTCGAGTTCATGCGCTGCGGTGCGCGCTTCTGCGCG 116
Db 28 GlyAlaPheProThrThrTy-----VallysLeuGlyPro-----Lys 40
QY 117 GGGNAGATG-----GAGGAGATCGCAACTTCCCGGTG 149
Db 41 GlyTyMetValTyArgProTyLeuLysAspAlaAlaAsnIleTyAsnMetProLeu 60
QY 150 CGGCCAGCAGCGGTGGATCGTCACCTACCCAGTCCGCGCACGAGTTCGTCACGAG 209
Db 61 ArgProThrAspValPheValAlaSerTyGlnArgSerGlyThrMetThrGlnGlu 80
QY 210 GTGCTCTACTTGTGGTACCGCAGGCGCT----- 236
Db 81 LeuValTrpLeuIleGluAsnAspLeuAsnPheGluAlaAlaLysThrTyMetSerLeu 100
QY 237 -----GACCCCGATGAGATCGGCTTCATGAAC 263
Db 101 ArgTyIleTyLeuAspGlyPheMetIleTyAspProGluLysGlnGluTyAsn 120
QY 264 -----ATCACGAGCAGCTCCCGGTCTCGAGTAC 293
Db 121 AspIleLeuProAsnProGluAsnLeuAspMetGluArgTyLeuGlyLeuLeuGluTy 140
QY 294 ---CCACAGCGCGCTGGACATCATCAAGAACTG-----ACCTCTCCCGCTCATC 344
Db 141 PheSerArgProGlySerSerLeuLeuAlaAlaValProThrGluLysArgPheVal 160
QY 345 AAGAGCCACCTGCGCTACCGCTTCTGCGCTCTACCTCCACATGGAGATCCCAAGTGC 404
Db 161 LysThrHisLeuProLeuSerLeuMetProProAsnMetLeuAsp---ThrValLysMet 179
QY 405 ATCTATATGGCTCGCAACCCCAAGGATCTGGTGTCTTATTATCATAGTTCACCGCTCT 464
Db 180 ValTyLeuAlaArgAspProArgAspValAlaValSerSerPheHisHisAlaArgLeu 199
QY 465 CTGCGGACCATAGCTACCGAGGACCTTTCAAGAACTTCTCGCGAGGATTTATGAATAT 524
Db 200 LeuTyLeuLeuAsnLysGlnSerAsnPheLysAspPheTrpGluMetPheHisArgGly 219
QY 525 AAGCTGGCTACGCTCTGTTGAGCAGTGGAGGTTCTGGGACGACCGCATGAGC 584
Db 220 LeuTyThrLeuThrProTyPheGluHisValLysValLysAlaLysArgHisAsp 239
QY 585 TCGAACGTCGCTTTTCTCAAGTATGAAGATCATCGGACCTCGGACCTGGTGGATGGAG 644
```

```

Db      240 ProhenMetLeuPheLeuPheTyrGluAspTyrLeuLysAspLeuProGlyCysIleAla 259
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy      645 CAGTGGCCAGATTCTCTGGGGTGTCTGTGACAAAGGCCAGCTGGAAGCCCTGACGGAG 704
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      260 ArgileAlaAppPheLeuGlyLysLysLeuSerGluGluGlnIleGlnArgLeuCysGlu 279
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy      705 CACTGCCACACAGCTGGTGACACAGCTGTCGAACGCTGAGGCGCTGCC----- 752
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      280 HisLeuAsn-----PheGluLysPheLysAsnAenGlyAlaValAsnMetGluAspTyr 297
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy      753 -----GTGGCCGGGGAAGAGTGGGGCTG 776
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      298 ArgGluIleGlyIleLeuAlaAspGlyGluHisPheIleArgLysGlyLysAlaGlyCys 317
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy      777 TGAAGGACATCTTCACCGCTCCATGAATGAGAAGTTTGACTTGTGTATATAAAGAGAAG 836
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      318 TrpArgAspTyrPheAspGluGluMetThrLysGlnAlaGluLysTrpIleLysAspAsn 337
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy      837 ATGGGAAGTGTGACCTCAGCTT 860
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      338 LeuLysAspThrAspLeuArgTyr 345
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 45
US-09-513-999C-6719
; Sequence 6719, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.Y.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59 US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 6719
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 12
; OTHER INFORMATION: Xaa=Glu or Gly

US-09-513-999C-6719

Alignment Scores:
Pred. No.:      3.18e-22      Length:      60
Score:          313.00      Matches:      58
Percent Similarity: 98.3%      Conservative: 1
Best Local Similarity: 96.7%      Mismatches: 1
Query Match:     7.1%      Indels:      0
DB:              2      Gaps:          0

US-10-768-158-1 (1-2419) x US-09-513-999C-6719 (1-60)

Qy      1467 ATGAGCAAAATTTCTACTAGGTCTTCAGAAATGGACGCTCTTCCAGAGACTTCCAG 1526
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1 MetSerArgAsnCysSerLeuGlyLeuGlnSer**ArgProSerAlaArgAspPheGln 20
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy      1527 CGGGGGCTCCAAAGGCCCAATGCAGAGGACCCCGGAGCATGTGCTGAGGGAGTCTG 1586
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      21 ArgAlaAlaProLysAlaGlnCysArgGlyAlaArgGlyAlaCysAlaGluGlySerLeu 40
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy      1587 CCTGGTGAAGCTGGCAGGTGGAGTCTAATGCAGTCAGGAGCATTTGTCAGTCAGTGGGTG 1646
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      41 ProGlyGluAlaGlyArgTrpGluSerAsnAlaValArgSerIleCysMetGlnTrpVal 60
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 46
```

```

US-09-795-926-2
; Sequence 2, Application US/09795926
; Patent No. 6555669
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING THE SAME
; FILE REFERENCE: LEX-0144-USA
; CURRENT APPLICATION NUMBER: US/09/795,926
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 303
; TYPE: PRT
; ORGANISM: homo sapiens

US-09-795-926-2

Alignment Scores:
Pred. No.:      7.18e-21      Length:      303
Score:          303.50      Matches:      74
Percent Similarity: 48.7%      Conservative: 61
Best Local Similarity: 26.7%      Mismatches: 99
Query Match:     6.9%      Indels:      43
DB:              2      Gaps:          9

US-10-768-158-1 (1-2419) x US-09-795-926-2 (1-303)

Qy      75 TACTTCGAGTTCCATGCGTGGCGCTGCG---CCCTTCTCGCGGGAAG---ATGGAG 128
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      27 PhePheThrTyrGlnGlyIleProTyrProIleThrMetCysThrSerGluThrPheGln 46
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy      129 GAGATCGCCAACTTCCCGGTGCGGCCCGACGCGTGGATCGTCACCTACCCCAAGTCC 188
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      47 AlaLeuAspThrPheGluAlaArgHisAspIleValLeuAlaSerTyrProLysCys 66
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy      189 GGCACCAGC-----TTGCTGAGGAGTGGTCTACTTGGTGAGCAGGGCGCT 236
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      67 GlySerAsnTrpIleLeuHisIleValSerGluIleTyrAlaValSerLys----- 84
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy      237 GACCCCGATGAGATCGCGCTTGATGAACATCGACGAGCAGCTCCCGGTCCTGGAGTACC 296
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      85 -----LysLysTyrLysTyrPro 90
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy      297 CAG---CCGGCGCTGGACATC-----ATCAAGGAATGACC 329
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      91 GluPheProValLeuGluCysGlyAspSerGluLysTyrGlnArgMetLysGlyPhePro 110
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy      330 TCTCCCGCGCTCATCAAGAGCCACCTGCCCTTACCGCTTTCTGCCCTCTGACCTCCACAAT 389
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      111 SerProArgIleLeuAlaThrHisLeuHisTyrAspLysLeuProGlySerIlePheGlu 130
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy      390 GGAGACTCCAGGTCACTATATATGGCTCGCAACCCCAAGGATCTGGTGTCTTATTAT 449
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      131 AsnLysAlaLysIleLeuValIlePheArgAsnProLysAspThrAlaValSerPheLeu 150
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```

Qy 450 CAGTTCACCGCTCTCTCGCGACCATGAGCTACCGAGGCACCTTTCAAGAAATTCGCGG 509
Db 151 HisPheHisAsnAspValProAspIleProSerTyrGlySerTrpAspGluPhePheArg 170
Qy 510 AGTTTATGATGATAAGCTGGCTACGGCTCCCTGGTTTGAGCAGCGTGACGAGTTCTGG 569
Db 171 GlnPheMetLysGlyGlnValSerTrpGlyArgTyrPheAspPheAlaIleAsnTrpAsn 190
Qy 570 GAGCAGCCGATGGATCGAAGCTGCTTTCTCAAGTATGAAGACATGCATCGGGACCTG 629
Db 191 LysHisLeuAspGlyAspAsnValLysPheIleLeuTyrGluAspLeuLysGluAsnLeu 210
Qy 630 GTGACGATGGTGAGCAGCTGCGCAGATTCCTGGGGGTGCTCTGTGACAAAGCCGACGTG 689
Db 211 AlaAlaGlyIleLysGlnIleAlaGluPheLeuGlyPheLeuThrGlyGluGlnIle 230
Qy 690 GAAGCCCTGACGGAGCACTGCACCAAGCTGTGTGACAGTGTGCAAGCTGAGGCCCTG 749
Db 231 GlnThrIleSerVal-----GlnSerThrPheGlnAlaMetArgAlaLysSerGln 247
Qy 750 -----CCGCTGGGC-----CGGGGAAGAGTTGGGCTGTGGGAG 782
Db 248 AspThrHisGlyAlaValGlyProPheLeuPheArgLysGlyGluValGlyAspTrpLys 267
Qy 783 GACATCTTCACCGCTCCCATGAATGAGAAGTTTCGACTTGGTGTATAAACAG 833
Db 268 AsnLeuPheSerGluIleGlnAsnGlnLeuMetAspGluLysPheLysGlu 284

RESULT 47
US-10-364-774-2
; Sequence 2, Application US/10364774
; Patent No. 6929937
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; FILE REFERENCE: POLYNUCLEOTIDES ENCODING THE SAME
; CURRENT APPLICATION NUMBER: US/10/364,774
; CURRENT FILING DATE: 2003-02-11
; PRIOR APPLICATION NUMBER: US/09/795,926
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 303
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-364-774-2

Alignment Scores:
Pred. No.: 7,18e-21 Length: 303
Score: 303.50 Matches: 74
Percent Similarity: 48.7% Conservative: 61
Best Local Similarity: 26.7% Mismatches: 99
Query Match: 6.9% Indels: 43
DB: 2 Gaps: 9

```

```

US-10-768-158-1 (1-2419) x US-10-364-774-2 (1-303)
Qy 75 TACTTTCAGTTCATGCGGTGGCGCTCGC---CCCTTCTCGCGCGGGAAG---ATGGAG 128
Db 27 PhePheThrTyrGlnGlyIleProTyrProIleThrMetCysThrSerGluThrPheGln 46
Qy 129 GAGATCCGCCAACTTCCCGGTGCGGCCGCCAGCGATGTGGATCGTCACCTACCCCAAGTCC 188
Db 47 AlaLeuAspThrPheGluAlaArgHisAspAspIleValLeuAlaSerTyrProLysCys 66
Qy 189 GGCACACAGC-----TTGCTCAGAGAGTGCTGTCTACTTGGTGAGCCAGGGCGCT 236
Db 67 GlySerAsnTrpIleLeuHisIleValSerGluLeuIleTyrAlaValSerLys----- 84
Qy 237 GACCCCGATGAGATCGCGCTTGATGAACATCGACGACGATCCCGGTCCTGGAGTACCCA 296
Db 85 -----LysLysTyrLysTyrPro 90
Qy 297 CAG---CCGGCGCTGGGACATC-----ATCAAGGAACTGACC 329
Db 91 GluPheProValLeuGluCysGlyAspSerGluLysTyrGlnArgMetLysGlyPhePro 110
Qy 330 TCTCCCGCTCATCAAGAGCCACCTGCTCCCTACCGCTTCTGCGCTCTGACCTCCACAAT 389
Db 111 SerProArgIleLeuAlaThrHisLeuHisTyrAspLysLeuProGlySerIlePheGlu 130
Qy 390 GGAGACTCCAGGTCTATATGGCTCGCAACCCCAAGGATCTGGTGTCTTATAT 449
Db 131 AsnLysAlaLysIleLeuValIlePheArgAsnProLysAspThrAlaValSerPheLeu 150
Qy 450 CAGTTCACCGCTCTCTCGCGACCATGAGCTACCGAGGCACCTTTCAAGAAATTCGCGG 509
Db 151 HisPheHisAsnAspValProAspIleProSerTyrGlySerTrpAspGluPheArg 170
Qy 510 AGTTTATGAATGATAAGCTGGCTACGGTCTCTGGTTTGAGCAGCTGCGAGGATTCGG 569
Db 171 GlnPheMetLysGlyGlnValSerTrpGlyArgTyrPheAspPheAlaIleAsnTrpAsn 190
Qy 570 GAGCAGCCGATGGACTCGAAGCTGCTTTTCTCAAGTATGAAGACATGCATCGGGACCTG 629
Db 191 LysHisLeuAspGlyAspAsnValLysPheIleLeuTyrGluAspLeuLysGluAsnLeu 210
Qy 630 GTGACGATGGTGAGCAGCTGGCCAGATTCCTGGGGGTGCTCTGCAAGGCCACGCTG 689
Db 211 AlaAlaGlyIleLysGlnIleAlaGluPheLeuGlyPhePheLeuThrGlyGluGlnIle 230
Qy 690 GAAGCCCTGACGGAGCACTGCGCACCAAGCTGTGTGAGCCAGTGTCTCAACGCTGAGGCCCTG 749
Db 231 GlnThrIleSerVal-----GlnSerThrPheGlnAlaMetArgAlaLysSerGln 247
Qy 750 -----CCGCTGGGC-----CGGGGAAGAGTTGGGCTGTGGGAG 782
Db 248 AspThrHisGlyAlaValGlyProPheLeuPheArgLysGlyGluValGlyAspTrpLys 267
Qy 783 GACATCTTCACCGCTCCCATGAATGAGAAGTTTCGACTTGGTGTATAAACAG 833
Db 268 AsnLeuPheSerGluIleGlnAsnGlnLeuMetAspGluLysPheLysGlu 284

RESULT 48
US-09-795-926-6
; Sequence 6, Application US/09795926
; Patent No. 6555669
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.

```



```
Db 52 ProGluPheProValLeuGluCysGlyAspSerGluLysTyTyrGlnAArgMetLysGlyPhe 71
Qy 327 ACCTCTCCCGCCTCATCAAGAGCAGCTGCCCTACCGCTTTCGCCCTCTGACCTCCAC 386
Db 72 ProSerProArgIleLeuAlaThrHisLeuHisTyrAspLysLeuProGlySerIlePhe 91
Qy 387 AATGAGACTCCAAAGGTCATCATATGCTCGCAACCCCAAGGATCTGGTGGTCTTAT 446
Db 92 GluAsnLysAlaLysIleLeuValIlePheArgAsnProLysAspThrAlaValSerPhe 111
Qy 447 TATCAGTTCCACCGCTCTCTCGGACCATGAGTACCGAGGACCTTTCAGAAATTCGC 506
Db 112 LeuHisPheHisAsnAspValProAspIleProSerTyGlySerTyAspGluPhePhe 131
Qy 507 CGGAGGTTTATGAATATAGCTGGGTACCGCTCCCTGGTTCGTGAGCAGCTGCAGGATTC 566
Db 132 ArgGlnPheMetLysGlyGlnValSerTrpGlyArgTyPheAspPheAlaIleAsnTrp 151
Qy 567 TGGGAGCAGCCGACTCGAAGCTGCTTTCTCAAGTATGAAGACATGCATCGGGAC 626
Db 152 AsnLysHisLeuAspGlyAspAsnValLysPheIleLeuTyGlnAspLysGluAsn 171
Qy 627 CTGTGTGAGTGTGGGACAGCTGCCAGATTCCTGGGGGTGTCTCTGTGCAAGGCCAG 686
Db 172 LeuAlaAlaGlyIleLysGlnIleAlaGluPheLeuGlyPhePheLeuThrGlyGluGln 191
Qy 687 CTGAAGCCCTGAGGAGCACTGCCACGAGCTGTGGACGAGTGTGCAAGCTGAGGCC 746
Db 192 IleGlnThrIleSerVal-----GlnSerThrPheGlnAlaMetArgAlaLysSer 208
Qy 747 CTG-----CCCGTGGGC-----CGGGAAGAGTTGGGCTGTGG 779
Db 209 GlnAspThrHisGlyAlaValGlyProPheLeuPheArgLysGlyGluValGlyAspTrp 228
Qy 780 AAGGACATCTTCACCGTCTCCATCAATGAGAAGTTTGACTTGGTGTATTAACAG 833
Db 229 LysAsnLeuPheSerGluIleGlnAsnGlnGluMetAspGluLysPheLysGlu 246

RESULT 50
US-09-854-122-20
; Sequence 20, Application US/09854122
; Patent No. 6841718
; GENERAL INFORMATION:
; APPLICANT: ALBERTE, RANDALL S.
; TITLE OF INVENTION: TRANSGENIC PLANTS INCORPORATING TRAIT OF ZOSTERA MARINA
; CURRENT APPLICATION NUMBER: US/09/854,122
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/202,529
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Flaveria bidentis
US-09-854-122-20

Alignment Scores:
Pred. No.: 4,32e-18 Length: 320
Score: 275.50 Matches: 80
Percent Similarity: 49.1% Conservative: 63
Best Local Similarity: 27.5% Mismatches: 111
Query Match: 6.3% Indels: 37
DB: 2 Gaps: 11

US-10-768-158-1 (1-2419) x US-09-854-122-20 (1-320)
Qy 72 AAGTACTTCAGATTCATCGGCTGGGTGGCTCGCGCCCTTCTCGCGGGGAAGATGAGGAG 131
Db 34 LysTyTyrGlnAspPheTrpGly-----LeuGlnAsnAsnIleGluGlyAlaIleLeuAla 51
```

```
Qy 132 ATCCGCAACTTCCGGTGGCGCCAGCGAGCTGTGGATCGTCACCTACCCCAAGTCCGGC 191
Db 52 GlnGlnSerPheLysAlaArgProAspAspValPheLeuCysSerTyTyrProLysSerGly 71
Qy 192 ACCAGCTGTCTGAGGAGGTGTCTACTTGTGTGAGCCAGCGGCTGACCCCGATGAGATC 251
Db 72 ThrThrTrpLeuLysAlaLeuAlaTyAlaIleValThrArgGluLysPheAspGluPhe 91
Qy 252 GGC-----TTGATGAACATCGAGGAGCAG-----CTCCCGCTCTCGGAGTACCCACAG 299
Db 92 ThrSerProLeuLeuThrAsnIleProHisAsnCysIleProTyTrpIleGlu----- 108
Qy 300 CCGGCGCTGACATCATCAAGGAACATGACCTCTCCCGC-----CTCATC 344
Db 109 -----LysAspLeuLysLysIleValGluAsnGlnAsnAsnSerCysPheThrProMet 126
Qy 345 AAGAGCCACCTGCCCTACCGCTTCTGCCCTCTGACCTCCACAATGGAGACTCCAAAGTC 404
Db 127 AlaThrHisMetProTyTrpHisValLeuProLysSerIleLeuAlaLeuAsnCysLysMet 146
Qy 405 ATCTATATGCTCGCAACCCCAAGGATCTGGTGTCTTATTAATCAGTTCCACCGCTCT 464
Db 147 ValTyTrpIleTyArgAsnIleLysAspValIleValSerPheTyTrpHisPheGlyArgGlu 166
Qy 465 CTGGGACCATGAGTACCGA---GGCACCTTCAAGAAATTCGCGGAGGTTTATGAAT 521
Db 167 IleThrLysLeuProLeuGluAspAlaProPheGluGluAlaPheAspGluPheTyHis 186
Qy 522 GATAAGCTGGGCTACCGCTCTCTGTGACGACGTGAGGAGTCTGGGAGCACCAGGATG 581
Db 187 GlyIleSerGlnPheGlyProTyTrpAspHisLeuLeuGlyTyTrpLysAlaSerLeu 206
Qy 582 GAC-----TCGACAGCTGCTTTCACAGTATGAAGACATGTCGCGGACCTGGTGACG 635
Db 207 GluArgProGluValIleLeuPheLeuLysTyTyrGluAspValLysLysAspProThrSer 226
Qy 636 ATGTGTGAGCAGCTGGCCAGATTCCTGGGGGTGTCTCTGT-----GACAAG 680
Db 227 AsnValLysArgLeuAlaGluPheIleGlyTyTrpPheThrPheGluGluLysGlu 246
Qy 681 GCCCAGCTGGAAGCCCTGACGAGCAGCTGC-----CACCAGCTGGTG 722
Db 247 GlyValIleGluSerIleIleLysLeuCysSerPheGluAsnLeuSerAsnLeuGluVal 266
Qy 723 GACCAGTGTGCAACGCTGAGCC---CTGCCCGTG-----GGCCGG 761
Db 267 AsnLysSerGlyAsnSerLysGlyPheLeuProIleGluAsnArgLeuTyTyrPheArgLys 286
Qy 762 GGAAGAGTGGGCTGTGGAAGGACATCTTCACCGTCTCCATGAATGAGAAGTTTGACTTG 821
Db 287 AlaLysAspGlyAspTrpLysAsnTyTrpPheThrAspGluMetThrGluLysIleAspLys 306
Qy 822 GTGTATAACAGAGATGGGAAAGTGTGACCTC 854
Db 307 LeuIleAspGluLysLeuSerAlaThrGlyLeu 317

RESULT 51
US-09-854-122-18
; Sequence 18, Application US/09854122
; Patent No. 6841718
; GENERAL INFORMATION:
; APPLICANT: ALBERTE, RANDALL S.
; APPLICANT: SMITH, ROBERT
; TITLE OF INVENTION: TRANSGENIC PLANTS INCORPORATING TRAIT OF ZOSTERA MARINA
; FILE REFERENCE: PHA-007.01
; CURRENT APPLICATION NUMBER: US/09/854,122
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/202,529
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 324
```

```
; TYPE: PRT
; ORGANISM: Brassica napus
US-09-854-122-18

Alignment Scores:
Pred. No.: 4.23e-17 Length: 324
Score: 265.50 Matches: 80
Percent Similarity: 47.3% Conservative: 61
Best Local Similarity: 26.8% Mismatches: 128
Query Match: 6.0% Indels: 29
DB: 2 Gaps: 11

US-10-768-158-1 (1-2419) x US-09-854-122-18 (1-324)
QY 45 CCCAGCACCCCGGGGAGTTCGAGAGCAAGTACTTCGAGTTCCATGGC---GTCCGGCTG 101
DB 28 ProSerGluLysGlyTrpLeuValCysGlnMetTyrGlnPheGlnGlyArgTrpHisThr 47
QY 102 CCGCCCTTCGCGCGGGAGAGTGGAGGATCGGCAACTTCCCGGTGGGCGCCAGCAGAC 161
DB 48 GlnAlaLeuLeuGlnGlyLeuThrCysGlnLysHisPheGluAlaLysAspSerAsp 67
QY 162 GTGTGGATCGTCACCTACCCCAAGTCGSCACCCAGCTTCTCGCAGGAGTGTCTAC--- 218
DB 68 llelleLeuValThrAsnProLysSerGlyThrTrpLeuLysAlaLeuValPheAla 87
QY 219 TTGGTGAGCCAGCGGCTGACCCCGATGAGATCGGCTTGATGAACATCGACGAGCTC 278
DB 88 LeuIleAsnArgHisLysPhePro---ValTyrSerValIlelleLeuSerCysTyrGln 106
QY 279 CCGGTCTCGGATACCCAGCCGGCTGGACATCATCAAG----- 320
DB 107 SerAlaLeuLeuValProPheLeuGlyArgSerLeuLeuArgSerProAspPheAspPhe 126
QY 321 ---GAATGACTCTCCCGCTCATCAAGACCCACTCGCTACCGCTTTCGCCCTCT 377
DB 127 SerGlnLeuSerSerProArgLeuMetAsnThrHisIleSerHisLeuSerLeuProGlu 146
QY 378 GACCTCCCAATGGAGACTCCCAAGGTCACTATATGGTTCGCAACCCCAAGGATCTGGTG 437
DB 147 SerValLysSerSerSerCysLysIleValTyrCysCysArgAsnProLysAspMetPhe 166
QY 438 GTGTCTTATATCAGTTCCACCGCTCTCTG-----CGGACCATGAGCTACCGAGGC 488
DB 167 ValSerLeuTrpHisPheGlyLysLysLeuAlaProGluGluThrAlaAspTyr----- 184
QY 489 ACCTTTCAAGATTCTCGCGAGGTTTATGATGATAGCTGGGCTAGGCTCTCGTTT 548
DB 185 ProIleGluLysAlaValGluAlaPheCysGlnGlyLysPheIleGlyLysProPheTrp 204
QY 549 GAGCAGCTGCAGAGTTCTGGGAGCACCGCATGGAC-----TCGAAGCTGTTTTTCTC 602
DB 205 AspHisValLeuGluTyrTrpTyrAlaSerLeuGluAsnProAsnLysValLeuPheVal 224
QY 603 AAGTATGAGACATGCATCGGACCTGTGGACGATGGTGAGCAGCTGGCCAGATCTCTG 662
DB 225 SerTyrGluGluProLysLysLysThrGlyGluThrIleLysArgIleAlaGluPheLeu 244
QY 663 GGGGTGTCC---TGTGACAGGCCAGCTGGAGGCTGAGGCTGACGAGCAGCTGC-----CAC 713
DB 245 GlyCysGlyLeuValGlyGluGluValArgAlaIleValLysLeuLysSerPheGlu 264
QY 714 CAGCTGGTGACAGCTGTCACGCTGAGGCC---CTGCCCGTGGGC----- 758
DB 265 SerLeuSerSerLeuGluValAsnArgGluGlyLysLeuProSerGlyMetGluThrArg 284
QY 759 -----CGGGGAAGAGTTGGCTGTGGAGGACATCTTCACCGCTCTCCATGAAT 806
DB 285 AlaPhePheArgLysGlyGluValGlyGlyTyrAspThrLeuThrGluSerLeuAla 304
QY 807 GAGAAGTTTGACTTGGTGTATAAACAGAGATGGGAAAGTGTGACTCACGTTT 860
DB 305 GluValIleAspArgThrIleGluGluLysPheGlnGlySerGlyLeuLysPhe 322
```

```
RESULT 52
US-09-854-122-19
; Sequence 19, Application US/09854122
; Patent No. 6841718
; GENERAL INFORMATION:
; APPLICANT: ALBERTE, RANDALL S.
; APPLICANT: SMITH, ROBERT
; TITLE OF INVENTION: TRANSGENIC PLANTS INCORPORATING TRAIT OF ZOSTERA MARINA
; FILE REFERENCE: PHA-007.01
; CURRENT APPLICATION NUMBER: US/09/854,122
; PRIORITY FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/202,529
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-854-122-19

Alignment Scores:
Pred. No.: 1.13e-16 Length: 302
Score: 261.00 Matches: 82
Percent Similarity: 44.7% Conservative: 53
Best Local Similarity: 27.2% Mismatches: 115
Query Match: 5.9% Indels: 52
DB: 2 Gaps: 13

US-10-768-158-1 (1-2419) x US-09-854-122-19 (1-302)
QY 6 GGCAGACGCGCGCGATGGCGAGAGCGAGCGAGCCAGACCCCGACGACCCCGC----- 56
DB 12 GlyAspGluAspLeuThrGlnGluThrArgAlaLeuIleSerSerLeuProLysGluLys 31
QY 57 GGGAGATTCGAGAGCAAGTACTTCGATTCCATGGCGTG---CGCTCGCGCCCTCTGCG 113
DB 32 GlyTrpLeuValSerGluIleTyrGluPheGlnGlyLeuTrpHisThrGlnAlaIleLeu 51
QY 114 CGCGGAAGATGGAGGAGATCGCAACTTCCCGGTGGCGCCAGCGAGCGTGTGGATCTGC 173
DB 52 GlnGlyIleLeuIleCysGlnLysArgPheGluAlaLysAspSerAspIleLeuVal 71
QY 174 ACCTACCCCAAGTCCGCGACCGAGCTGCTGCAGGAGGTGGTCTACTCTGGTG----- 224
DB 72 ThrAsnProLysSerGlyThrTrpLeuLysAlaLeuValPheAlaLeuAsnArg 91
QY 225 -----AGCCAGGCGCTGACCCCGATGAGATCGGCTTGATGAACATC 266
DB 92 HisLysPheProValSerSerSerGlyAsnHisPro-----LeuLeuValThrAsnPro 109
QY 267 GACGAGCAGCTCCCGGTCTCGAG-----TACCCACAGCGCGGCTCGACATCATC 317
DB 110 HisLeuLeuValProPheLeuGluGlyValTyrTyrGluSerProAspPheAsp---Phe 128
QY 318 AAGGAACCTGACTCTCCCGCTCATCAAGAGCCACCTACCGCTTCTGCGCTCT 377
DB 129 SerSerLeuProSerProArgLeuMetAsnThrHisIleSerHisLeuSerLeuProGlu 148
QY 378 GACCTCCCAATGGAGACTCCCAAGGTCTATATGGTTCGCAACCCCAAGATCTGTG 437
DB 149 SerValLysSerSerSerCysLysIleValTyrCysCysArgAsnProLysAspMetPhe 168
QY 438 GTGTCTTATTATCAGTTCCACCGCTCTCTG-----CGGACCATGAGTACCGAGGC 488
DB 169 ValSerLeuTrpHisPheGlyLysLeuAlaProGluGluThrAlaAspTyr----- 186
QY 489 ACCTTTCAAGAACTTCTCGCGGAGGTTTATGAATGATGAAGTGGGTACCGGCTCTCGTTT 548
DB 187 ProIleGluLysAlaValGluAlaPheCysGluGlyLysPheIleGlyLysProPheTrp 206
QY 549 GAGCAGCTGCGAGGATCTCTGG-----GAGCACCAGCATGGACTCGAACGCTGCTTTTCTC 602
```



```
Alignment Scores:
Pred. No.: 1.21e-11 Length: 636
Score: 212.00 Matches: 96
Percent Similarity: 36.8% Conservative: 11
Best Local Similarity: 33.0% Mismatches: 96
Query Match: 4.8% Indels: 88
DB: 2 Gaps: 20

US-10-768-158-1 (1-2419) x US-09-252-991A-32429 (1-636)
QY 798 AGAGCGTGAAGATGCTCTCCACAGCCCAACTCTTCCCGCGCCACGGGAGGCGCTCAG 739
Db 1 ArgArgGlyArgCysSerGlyArgLeu-----ArgArgGlyProGly 14
QY 738 CGTTGCAGCACTGCTCCACAGCTGTGGCAGTGTCTCCGTCAGGCGCTTCCAGCTGGGCT 679
Db 15 ArgGlySer-----ProGlyThrAlaArgSerAlaProAlaGlyAspArgGly--- 31
QY 678 TGTCAAGACACACCCAGGAAATCTGG-----CCAGCTGTCCACCATCTGTC 631
Db 32 ---ArgArgArgArgProGlyThrTrpArgArgArgLeuProGlySerAlaProAla 50
QY 630 CCAGGTCCTCGATGCTTTCATCTTGAGAAAGACACGTTCCAGTCCATCGGCTGT 571
Db 51 ProArgProAspAla----- 55
QY 570 CCAGAACTCTCTGCACGCTGCTCAAAACAGGAGCGGTAGCCAGCTTATCATTTATAAAC 511
Db 56 ---ArgArgProGlyCysAlaGlyArgArgGlyArgSer-----Arg 69
QY 510 TCGGCGAATTTCTGAAAGGTGCTCGGTAGTCA-----TGG----- 472
Db 70 ProGlyArgArgGlyArgCysSerAlaGlySerArgProGlyTrpArgTrpPro 89
QY 471 TCCGACAGAGCGGT-----GGAAGTGTATAAGACACCA 433
Db 90 AlaAlaThrAlaGlyArgCysAlaGlyArgCysGlyArgSerArgThrAlaPro 109
QY 432 GATCTTGGGTTTCGAGAGCATATAGATGACCTTGGAGTCTCCATTTGGAGGTCAG--- 376
Db 110 SerSerIleGly-----GlyTrpArg---HisAlaGlyGlyAspGly 122
QY 375 AGGCAGAAAGCGGTAGGCGAGTGGCTTGTATGAGCGGGGAGAGTCAAGTTCCTTGA 316
Db 123 ArgArgArgArgGlyArg-----GlyHisArgAspAsnProCys 135
QY 315 TGATGTCAGGCCCG-----GCTGTGGTACTCCAGGACCGGAGTGTCTCGTGTGT 262
Db 136 ArgCysSerArgAspTrpArgArgAspGlyArgProGlyAlaAlaGlyArgArgThrCys 155
QY 261 TCATCAAGCGCATCTCATCGGGTTCAGCGCTCGCTCACCAGTAGACCACTCTCTGCA 202
Db 156 ArgSer-----ArgProGly-----ProProGlyPro 164
QY 201 GCAAGCTGTGTCGG-----ACTTGGGTAGGTGAGCATCCACAGTCTCGTGGGCC 151
Db 165 TrpHisTrpProArgArgGlyArgAlaGlyArgGlyArgGlySerAlaArgTrpLeu 184
QY 150 GCACCGGAAGTTGGCGATCTCTCCATCTTCCCGCGGAGAGGCGGAGCGGCGCAGC 91
Db 185 ProProGlySerAlaArgSerValAlaAlaArgArgGly---ArgAlaAsnArgCysArg 203
QY 90 CATGGAAGTCTGAAGTACTTGTCTCGAACTCCCCCGGGGTGTGTGG---GGGTCTCTCGGCT 34
Db 204 ArgGly-----ArgArgProProGlyArgArgProGlyArgArgCys 217
QY 33 CGCTCTCCGCATGCCCGCCCGCTGCGCCGTCG 1
Db 218 ArgThrArgProAlaArgArgArgArgArg 228

RESULT 56
US-09-252-991A-25394
; Sequence 25394, Application US/09252991A
```

```
Alignment Scores:
Pred. No.: 4.52e-16 Length: 214
Score: 254.00 Matches: 55
Percent Similarity: 53.3% Conservative: 42
Best Local Similarity: 20.2% Mismatches: 73
Query Match: 5.8% Indels: 12
DB: 2 Gaps: 3

US-10-768-158-1 (1-2419) x US-10-364-774-16 (1-214)
QY 315 ATCAGGAGTACGCTCTCCCGCTCATCAGAGCCACCTCCCTACCGCTTCTGCGCC 374
Db 17 MetLysGlyPheProSerProArgIleLeuAlaThrHisLeuHisTyrAspLysLeuPro 36
QY 375 TCTGACCTCCCAATGGAGCTCCAAGGTCTATATATGGTTCGCAACCCCAAGGATCTG 434
Db 37 GlySerIlePheGluAsnLysAlaLysIleLeuValIlePheArgAsnProLysAspThr 56
QY 435 GTGTGTCTTATTATCACTCCACGCTCTCTCGGACCATGAGTACCGAGGACCTTT 494
Db 57 AlaValSerPheLeuHisPheHisAsnAspValProAspIleProSerTyrGlySerTrp 76
QY 495 CAAGAATTCTCCGAGGTTTATGATATGAATGATGAAGCTGGGCTACGGCTCTGTTGAGCAC 554
Db 77 AspGluPhePheArgGlnPheMetLysGlyGlnValSerTrpGlyArgTyrPheAspPhe 96
QY 555 GTGAGGAGTTCGGAGCACCGCATGACTCGAACGCTGCTTTTCTCAAGTATGAAGAC 614
Db 97 AlaIleAsnTrpAsnLysHisLeuAspGlyAspAsnValLysPheIleLeuTyrGluAsp 116
QY 615 ATGCATCGGACCTGGTGACGATGTGGAGCAGCTGGCCAGATTCCTGGGGGTCTCTGT 674
Db 117 LeuLysGluAsnLeuAlaAlaGlyIleLysGlnIleAlaGluPheLeuGlyPhePheLeu 136
QY 675 GACAAGCCAGCTGGAAGCCCTCAGGAGCAGTGCACGAGCTGGTGAGCAGTGTGTC 734
Db 137 ThrGlyGluGlnIleGlnThrIleSerVal-----GlnSerThrPheGlnAlaMet 153
QY 735 AACGCTGAGGCGCTG-----CCCTGGGC-----CGGGGAAGA 767
Db 154 ArgAlaLysSerGlnAspThrHisGlyAlaValGlyProPheLeuPheArgLysGlyGlu 173
QY 768 GTTGGGTGTGGAAGGACATCTTCAACGCTCTCCATGATGAAGAAAGTTTGACTTGTGTAT 827
Db 174 ValGlyAspTrpLysAsnLeuPheSerGluIleGlnAsnGlnGluMetAspGluLysPhe 193
QY 828 AACAG 833
Db 194 LysGlu 195

RESULT 55
US-09-252-991A-32429
; Sequence 32429, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32429
; LENGTH: 636
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32429
```

[illegible]

[illegible]

```

RESULT 58
US-09-252-991A-26169
; Sequence 26169, Application US/09252991A
; Patent NO. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26169
; LENGTH: 282
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26169

```

Alignment Scores:			
Pred. No.:	1.11e-09	Length:	282
Score:	190.00	Matches:	89
Percent Similarity:	37.1%	Conservative:	23
Best Local Similarity:	29.5%	Mismatches:	106
Query Match:	4.3%	Indels:	84
DB:	2	Gaps:	16
US-10-768-158-1 (1-2419) x US-09-252-991A-26169 (1-282)			
QY	816	CAAACTCTCTCATTCATGAGACGGTGAAGATGTCCTTCCACAGCCCACTCTTCCCGCGC	757
DB	10	ArgrHmetArGlyArGArGArGArGProCysAlaGlyCysAlaArGAlaArGArgGly	29
QY	756	CCAGGGCAGGCGCTCAGCGTTGACAGACTGGTCCACAGCTGGTGGCAGTGTCCGTCA	697
DB	30	ProAlaSerArGProArGlyCysSerAlaGlyArGProAlaProGlyGlySerProSer	49
QY	696	GGGCTTCCAGCTGGG-----CCCTGTCTCAGGACA	667
DB	50	AlaSerArGHisGlyArGAlaTrpArGArGSerProAlaProProCysSerAlaArG	69
QY	666	CCC-----CCAGGAATCTGGCCAGCT-----GCTCCACCATCGTCA	631
DB	70	ProGlyProAlaArGlyProGlyArGArGProGlyArGlyArGThrProlySerAla	89
QY	630	CCAGTCCCGATGTCATGCTTTCATACTTTCAGAAAAAGCA---CGTTCGAGTCCATCGCGT	574
DB	90	ProValSerProAlaProArGHisArGProArGlySerAlaArGArGAlaSerProArGln	109
QY	573	GCTCCAGAACTCTTCGACGTGCT-----CAACACGAGCGCTAGCCAGCTTATCAT	520
DB	110	GlyProAlaArGProSerArGAlaAsnArGSerAlaAlaSerArGArGProAlaAlaGly	129
QY	519	TCATAAACCTCCGCGAGAACTTTCGAAGGTGCCTCGGTAGTTCATGTCGCCAGAGAC	460
DB	130	AlaGlyThrGlyAlaArGThrCysSerArG-----	139
QY	459	GGTGGAACTGATAATAAGACACCACAGATCCTTGGGGT-----ProAlaProGlyGlyArGSerArGThrLeuArGAla	421
DB	140	-----ProAlaProGlyGlyArGSerArGThrLeuArGAla	151
QY	420	TGCAGACCATATAGATGACCTTGGAGTCTCATTTGTGGAGTTCAGAGGCAGAAAGCGGT	361
DB	152	SerAlaProHisArGSerAlaArGProArGTrpThrProGlyThrArg---LysAlaGly	170
QY	360	AGGCGAGT---GGCTCTTGATAGCGCGGAGAGAGTTCCTTCGTATGATGATCCAGGC	304
DB	171	ArgSerGlyGlnGlyLeuArGProGlyAlaSerArGysGly-----CysAlaGly	187
QY	303	CCGGCTGTGGGTACTCCAGGACCGGGAGCTGCTCGATGTTTCATCAAGCCGATCTCAT	244
DB	188	GlySerAlaThrThrProAlaProArGSerVal-----	198
QY	243	CGGGGTACGCGCCCTGGCTCACCAAGTAGA---CCACCTCTTCGACCAAGCTGGTCCCGG	187
DB	199	-----ProProSerProAlaArGTYrProAlaProValAlaGly-----	211
QY	186	ACTTGGGGTAGTGACATCACAGTCTCGTGGGCGGCACCGGGAAGTGGCGATCTCT	127
DB	212	-----ArgAspArGArGSerProProAlaProGlySer-----	222
QY	126	CCATCTTCCCGGCGAGAGGGCGCAGCCGACGCCCATGGAACCTCGAAGTACTTGTCTCT	67
DB	223	ProAlaSerArG---ArgLeuSerArGThrAspArGProGlyAsnArGAla-----	238
QY	66	CGAACTCCCGGGGTGCTGGGGGTCTCGGCCTCGCTCTCCGCCATGCGCGCGCGTCCG	7
DB	239	-----ProSerProAlaGlyAsnArGGlutlleHisProAlaCysGlnProLeuPhe	255
QY	6	CGTGGC 1	
DB	256	ArgArg 257	

QY 243 GATGAG---ATCGGCTTATGAACATCGACGAGCGCTCCGGTCTCGAGTAC----- 293
Db 123 ThrHisProLeuLeuPhePheAsnProHisSerCysValGlnAenLeuGluLeuTyr 142
QY 294 -----CCACAGCCGGCTGACATCATCAAGGAACCTCCCGC 338
Db 143 MetGlyArgGluAenThrMetProAspLeuAspMetLeuAenGlu-----SerProArg 160
QY 339 CTCATCAAGACCCACCTGCTCCCTACCGCTTCTGCTCTGACCTCCCAATGAGAGATCC 398
Db 161 LeuPheAlaGlyHisIleProTyr-SerLeuLeuProAlaSerValLeuLysSerGlyThr 180
QY 399 AAGTCTATATGGCTCGCAACCCCAAGGATCTGGTGTCTTATATCACTCCAC 458
Db 181 LysIleIleAenIleSerArgAsnArgLysSerThrPheValSerPheTrpLysPheGly 200
QY 459 CGCTCTCTGCGGACC----- 473
Db 201 AsnLeuIleAenProAspLysLeuLeuAspLysLeuLysSerValAspIlePheAlaSer 220
QY 474 ---ATGAGTACCGAGCACCTTTCAAGAAATCTGCGGAGGTTTATGAATGATAAGCTG 530
Db 221 GlyIleSerPheCysGlyProGluTrpAenPheGlnAlaGluPheThrAsn----- 237
QY 531 GGCTACGGCTCTGGTTTGACACGTCGAGAGTTCTGGGAGCACCGCATCGACTCGAAC 590
Db 238 -----AlaIleSerThrAsnSerAsn 244
QY 591 GTGCTTTTCAAGTATGAAGACATGATCGGACCTGTGACGATGTGGAGCAGCTG 650
Db 245 LeuLeuLeuLeuSer-TyrGluGluMetLeuLysProValGluAenValLysLysLeu 264
QY 651 GCCAGATTCCTGGGGTGTCC---TGTGACAAAGCCCGCAGCTGGAAGCCCTGACGGAGCAC 707
Db 265 AlaGluPheMetGlyCysGlyPheThrAspAspGluGluLysGlnGlyIleValAspGlu 284
QY 708 TGCCACCAAGCTG-----GTGACACAGTGTGCAACGCTGAGGCCCTGCCCGTGGC--- 758
Db 285 IleValLysLeuCysSerPheAspAsnLeuLysAenGlnValAsnLysAenGlySer 304
QY 759 -----CGGGGAAGAGTTGGGCTGTGG 779
Db 305 SerTyrAenSerLysIleAspAsnLysHisPheArgLysGlyGluValArgAspTrp 324
QY 780 AAGGACATCTCACGCTCCCATGATGAGAGTTTGAATGTTGGTATATAA 830
Db 325 AlaAenTyrLeuThrSerGluMetIleLysLysLeuGluThrAlaGlyLys 341

RESULT 63
US-09-252-991A-25346
; Sequence 25346, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25346
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25346

Alignment Scores:
Pred. No.: 3.63e-09 Length: 252

Score: 184.50 Matches: 89
Percent Similarity: 35.4% Conservative: 16
Best Local Similarity: 30.0% Mismatches: 98
Query Match: 4.2% Indels: 94
DB: 2 Gaps: 16

US-10-768-158-1 (1-2419) x US-09-252-991A-25346 (1-252)

QY 822 CCAAGTCAAATCTTCTATTTCATGGAGACGGTGAAGATGTCTCTCCACAGCCCACTCTTC 763
Db 21 ProGlyGlnArgSerCysArgTrpArg-----ProProProAlaSerAl 36
QY 762 CCGCGCCACCGGGCAGGCGC-----TCAGCCTTGCCAGCACTGGT 724
Db 36 a-ArgProProArgArgProArgTyrValAlaProPheHisSerAlaAlaGlnHisAlaS 56
QY 723 CCACCACTGGTGGCAGTGTCCGCTCAGGGCTTCCAGCTGGGCTTCTCAGAGACACC 664
Db 56 erAla-ProGlnGlnLeuAlaValAlaThrProCysAspProAlaCysArgSerAsp 75
QY 663 CCAGGAATCTGGCAGCTGTCCACCATCTCCACCATCTCCAGGTCCTCCGATGCTCTTCATACT 604
Db 76 SerAlaIleSerAlaProAlaProAlaProAlaPro-----CysIle----- 91
QY 603 TGAGAAAAGCACGTTCCAGTCCATGCGGTGCTCC-----AGAACTCTCTGCACTG 553
Db 92 -----ProAlaGlySerProProGlyAlaArgSerGlyLeuArg 104
QY 552 GCTCAAAACAGGAGCCGTAGCCAGCTTATCATTTCAATAAACCTCCGGCAGAAATCTTGA 493
Db 105 AlaSerArgArgSerArgSerProAla----- 113
QY 492 AGTGCTCTCGTAGCTCATGTCGCGACAGCGGTGGAACGTGATAATAAGACACCA 433
Db 114 -----ProArgSerArgProThrAlaCysArgProProPro 125
QY 432 GATCCTTTGGGTTGCGAGCCATATAGATGACCTGGAGTCTCCATTTGAGAGTCAAGG 373
Db 126 AlaPro-----ProProArgAlaProThrAspAlaProAlaGlyAlaProArg 141
QY 372 GCAGAAACGGTAGGGCAGGTGCTTTGTATGAGCGGGGAGAGTCACTTCTTGATGA 313
Db 142 SerArgProAlaArg----- 146
QY 312 TGTCAGGCGCGGTGGTGTCTCCAGGACCGGAGCTGCTCGTGTTCATCAAGC 253
Db 147 CysAlaGlyArgAla---AlaSerProGlyArgSerAlaArgArgThrArgSerPro 165
QY 252 CGATCTCATCGGGTCAAGCGCTTGGC-----TCACCAAGTAGACCACTCTCTGCA 202
Db 166 Arg---ArgArgGlySerAlaProGlyArgArgCysGlyThrSerAlaProGlnProArg 184
QY 201 GCAAGCTGGTCCCGACTTGGGTGAGTACAGATCCACAGTCCGCTGGGCGCGACCGGGA 142
Db 185 ThrGlyTrp-----ArgArgAsnAlaArgGlnSerAlaProAsp 197
QY 141 AGTTGGCGATCT-----CCTCCATCTTCCCGCGGAGAGGCG 103
Db 198 AlaGlyGlySerHisGlyProGlyProAspArgProGlyThrAlaArgArgAlaArgPro 217
QY 102 GCAGCCGACGCCATGGAACCTCGAAGTACTTGTCTCGAACTCCCGCGGGTGTGGGGG 43
Db 218 ProValProAlaProGlyCysHisArgGlyCysSerArg-----ProGly---TrpArg 234
QY 42 TCT-----CGGCTCTGCTCTCCGCATGCGCG 16
Db 235 SerAlaGlyArgSerAlaProArgProCysProProCysArg 249

RESULT 64
US-09-949-016-9748
; Sequence 9748, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:


```
Db 250 -----AlaSerArgProSerSerAlaAlaCysProMetAlaArgProThrThrTrp--- 266
Qy 694 CCCTGACGGAGCAGCTGCCACCACTGGTGACCACTGCTGCAACGCTGAGGCCCTGCCCG 753
Db 267 -----SerThrAlaThrGly---TrpThrSerThrProAlaAlaSerAlaTrpPro 282
Qy 754 TGGCGCGGGAGAGTTGGCTGT-----GGAAGACATCTTCACGG 795
Db 283 TrpAlaArgProProAlaAlaCysGlySerAlaArgThrAlaGlyArgThrGlySerSer 302
Qy 796 TCTCCA 801
Db 303 SerPro 304

RESULT 67
US-09-252-991A-28625
; Sequence 28625, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28625
; LENGTH: 202
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28625

Alignment Scores:
Pred. No.: 1,96e-08 Length: 202
Score: 176.50 Matches: 77
Percent Similarity: 30.8% Conservative: 9
Best Local Similarity: 27.6% Mismatches: 86
Query Match: 4.0% Indels: 107
DB: 2 Gaps: 14

US-10-768-158-1 (1-2419) x US-09-252-991A-28625 (1-202)
:::|:::|
825 ACACCAAGTCAAATCTTCATTCTGAGAGCGGTGAAGATGCTCTCCACAGCCCAACTC 766
6 SerProThrAlaProProArgArgTrpAlaArg----- 16

Qy 765 TTCCCGCGCCACCGGGCAGGGCCTCAGCGTTGTCAGCACTGGTGTCCACAGCTGGTGGCAGT 706
Db 17 -----ProArgArgSerArgThrArgCysAlaProProArgAlaGlyArgPro 33
Qy 705 GCTCGTCAAGGCTTCCAGCTGGGCTTGTTCACAGACACCCCGAGGAATCTGGCCA--- 649
Db 34 AlaArgAlaAlaProProAlaGlyArgArgArgSerPro-----CysTrpProArg 51
Qy 648 GCTGCTCCACCATCTGTACCAGTCCCGATCCGATGCTTCTCATCTTGAGAAAAACCACT 589
Db 52 ArgCysArgArgArgThrProGlyProProArgCys-----AlaGly 65
Qy 588 TCGAGTCCATGGGTGCTCCAGAACTCCTGCACGTGCTCAACACGAGGC----- 538
Db 66 SerArgProSerGlyArgProArgArgArgSerArgTyrAlaArgSerTrpProGly 85
Qy 537 -----CGTAGCCCAAGCTTATCATTTATAAACCTCCGGCAGAAATCTTTGAAAGGTGC 487
Db 86 GlyLeuArgArgHisProAla----- 92
Qy 486 CTCGGTAGCTCATGTGTCGCAGAGAGCGGTGGAACCTGATATAAGACACCACAGATCCT 427
Db 93 -----ArgGlyGlySerGlyCysLeuArgGlyThrAlaArg 104
```

```
Qy 426 TGGGGTTCGAGCCCATATAGATGACCTTGGAGTCTCCATTGTGGAGGTTCAGAGGCGAGAA 367
Db 105 TrpGlyGlyArgProValArg-----ArgAlaSerGlyArgArgProPro 119
Qy 366 ---AGCGGTAGCGAGGTGGCTCTTGATGAGCGGGAGAGGTCAAGTTCCTTGATGATGT 310
Db 120 GlySerGlyArgArgAspGlySer-----GlyProGlyArgArgSerPro----- 134
Qy 309 CCAGGCCCGGTGGGTACTCCAGAGCGGGAGTGTCTCGATGTTTCATCAAGCCGA 250
Db 135 -----CysGlyValProGly----- 139
Qy 249 TCTCATCGGGGTGAGCGCCCTGGCTCAACCAAGTAGACCACTCTCTGCAGCAAGCTGGTGC 190
Db 140 -----ArgProGlySerPro----- 144
Qy 189 CGGACTTGGGTAGGTGACGATCCACAGCTCGCTGGGCGGCACCG-----CGA 142
Db 145 -----ArgArgTrpArgArgProValArgArgThrGly 155
Qy 141 AGTTGGCGATCTCCTCCA-----TCTTCCCGCGCAGAGGGCGGCAGCC 97
Db 156 SerArgArgGluSerProArgSerValArgArgProAlaGlyArgArgAlaValPro 175
Qy 96 GCAGCCCATGAACTCGAAGTACTTGTCTCGAACTCCCGCGGGGTGCTGGGGGTCT 40
Db 176 AlaTrpArgGlyLeuValLeuHisCysSerCysHisValProHisAlaTrpGlyThr 194

RESULT 68
US-09-252-991A-23981
; Sequence 23981, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23981
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23981

Alignment Scores:
Pred. No.: 3,45e-08 Length: 525
Score: 176.50 Matches: 96
Percent Similarity: 34.2% Conservative: 14
Best Local Similarity: 29.8% Mismatches: 107
Query Match: 4.0% Indels: 105
DB: 2 Gaps: 20

US-10-768-158-1 (1-2419) x US-09-252-991A-23981 (1-525)
Qy 786 TGTCTCTTCCACAGCCCAACTCTTCCCGCGCCAGGGCGGCTTCAGCGTTGCAGCACT 727
Db 11 CysPro-----GlnProSerProGlyArgArgArgSerProLeuArgGlnArgArg 27
Qy 726 GGTCCACAGCTGTGGCAGTGTCCGTCA-----GGGCTTCCAGCTGGGCCT 679
Db 28 GlyAlaPro-----GlyGlnArgProAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 45
Qy 678 TGTACAGGACACCCCGAGGATCTGGCCAGTCTCCACCATCTCCACAGGTCCCGAT 619
Db 46 ArgGlyAlaArgGluGlnGly-----ProArgAlaGlnArgProAla 59
```

```
Qy 618 GCATGCTCTCATCTTGAGAAAGACGCTTCGAGTCATCGGTGCTCCAGCACTCCT 559
Db 60 GlnCysArgPheArgLeuProArgAlaAlaHisHisProGlnProArgPro 79
Qy 558 GCACGT----- 553
Db 80 ThrGlnGlyArgArgSerLeuArgProGlyHisArgThrArgHisProArgArgGlnArg 99
Qy 552 -----GCTCAACACAGAGCCGTAGCCAGCTTATCATTTCAATAACC----- 511
Db 100 ProValAlaArgHisArgProArgProGlyValProArgArgThrGlyProValArg 119
Qy 510 -----TCCGGCAGATCTTGAAGGTGCTCCGTAGCTCATGTCCGCA 466
Db 120 GlyAspProAlaSerAlaArgArgAlaGlyArgAlaGlyAlaArgArgGlyAla 139
Qy 465 GAGAGCGGTGAATCTGATAAAGACACACCAGATCTCTGGGTTC-----GAG 415
Db 140 ArgSerGlyThrGluGlyLys-----CysArgArgGlyGln 152
Qy 414 CCATATAGATGACCTTGGAGTCTCCATTTGGAGGTGAGAGTGCAGAGCGGAGCGGTAGGCA 355
Db 153 Pro-----GlyGlnArgAlaAspGlyValArgArg 162
Qy 354 GGTGGCTTTGATGAGCGGGGAGAGGTCACTTCTTCATGATGTCAGGCGCCGCTGTG 295
Db 163 GlyProPro-----AlaGlyAspArgArgThrPro-----LeuArgProGlyProAlaAla 179
Qy 294 GGTACTCCAGACCG---GGAGCTGCTCGTCGATGTTTCATCAAGCCGATCTCATCGGGGT 238
Db 180 AlaLeuProGlyProArgProAlaProArgAlaLeuProLeuSerArgProGlyArgGly 199
Qy 237 CAGCGCTGCTGCTCACCAGTAGACCACTCTCTGAGCAAGCTGGTCCGCGACTGGGGT 178
Db 200 ProGlyProGlyArgArgGlnAlaArgProAlaGly-----Gly 212
Qy 177 AGGTGACATCCACACCTGCTGGCGCCGACCG-----GGAAGTTGG 136
Db 213 Arg-----ArgArgArgAlaGlnProValAlaGlnArgProAlaGlyHisArg 228
Qy 135 CGATCTCTCCATCTTTC---CGCGCAGAGGCGCGCAGCGCATGGAACCTCGA 79
Db 229 GlnAspProProGlyGlnProProAlaArgProAlaAlaGlyAlaArgArgGly---Arg 247
Qy 78 AGTACTGCTCTCGAATCTCCCGGGGTGCTGGGGGTCTCGGCTCTGCTCTCCG----- 25
Db 248 SerProGlyGlyArgGlyAspPro-----PheGlyGlyGlnProArgProGlnAla 265
Qy 24 -----CCATCGCGCGCGCGT-----CGC 7
Db 266 LeuAlaAlaAlaThrValProProAlaAlaProLeuArgLeuArgAlaGlyProGlyArg 285
Qy 6 CGTGGC 1
Db 286 ArgArg 287

RESULT 69
US-09-795-926-8
; Sequence 8, Application US/09795926
; Patent No. 6555669
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Waite, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Potter, David George
```

```
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING THE SAME
; FILE REFERENCE: LEX-0144-USA
; CURRENT APPLICATION NUMBER: US/09/795,926
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 148
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-795-926-8

Alignment Scores:
Pred. No.: 1,83e-08 Length: 148
Score: 176.00 Matches: 42
Percent Similarity: 47.7% Conservative: 31
Best Local Similarity: 27.5% Mismatches: 50
Query Match: 4.0% Indels: 30
DB: Gaps: 5

US-10-768-158-1 (1-2419) x US-09-795-926-8 (1-148)
Qy 126 GAGGAGATCGCAACTTCCCGGTGCGGCCAGCCAGCATCGTGGATCGTCACCTACCCCAAG 185
Db 8 GlnAlaLeuAspThrPheGluAlaArgHisAspAspIleValLeuAlaSerTyrProLys 27
Qy 186 TCAGGACACAGC-----TTGCTGAGAGGTGCTACTTGGTGGAGCCAGGC 233
Db 28 CysGlySerAsnTrpIleLeuHisIleValSerGluLeuIleTyrAlaValSerLys--- 46
Qy 234 GCTACCCCGGATGAGTCGGCTTCATGAACATCCAGCAGCATCCCGGCTCTCGAGTAC 293
Db 47 -----LysLysTyrLysTyr 51
Qy 294 CCACAG---CGGCGCTGGACATC-----ATCAAGGAACCTG 326
Db 52 ProGluPheProValLeuGluCysGlyAspSerGluLysTyrGlnArgMetLysGlyPhe 71
Qy 327 ACCTTCTCCCGCTCATCAAGAGCCACTGCGCTTACCGCTTCTGCGCTCTGACCTCCAC 386
Db 72 ProSerProArgIleLeuAlaThrHisLeuHisTyrAspLysLeuProGlySerIlePhe 91
Qy 387 AATGAGACTCCAGGTCACTATATGCTGCGCAACCCCAAGGATCTGGTGTCTTAT 446
Db 92 GluAsnLysAlaLysIleLeuValIlePheArgAsnProLysAspThrAlaValSerPhe 111
Qy 447 TATCAGTTCCACCGCTCTCTCGCGGACCATGAGTACCGAGGACCTTTTCAAGAATTCGCG 506
Db 112 LeuHisPheHisAsnAspValProAspIleProSerTyrGlySerTrpAspGluPhePhe 131
Qy 507 CGGAGGTTTATGAATGATAAGCTGGGTACGGCTCTGCGTCTGCG 545
Db 132 ArgGlnPheMetLysGlyGlnGluSer---GlyCysTrp 143

RESULT 70
US-10-364-774-8
; Sequence 8, Application US/10364774
; Patent No. 6929937
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
```

```

; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING THE SAME
; FILE REFERENCE: LEX-0144-USA
; CURRENT APPLICATION NUMBER: US/10/364,774
; CURRENT FILING DATE: 2003-02-11
; PRIOR APPLICATION NUMBER: US/09/795,926
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 148
; TYPE: PRT
; ORGANISM: homo sapiens
; IS-10-164-774-8

```

Alignment Scores:		
Pred. No.:	1.83e-08	148
Score:	176.00	42
Percent Similarity:	47.7%	31
Best Local Similarity:	27.5%	50
Query Match:	4.0%	30
DB:	2	5
		Gaps:
		Length:
		Matches:
		Conservative:
		Mismatches:
		Indels:
		Gaps:

US-10-768-158-1 (1-2419) x US-10-364-774-8 (1-148)

126	QY	GAGGAGATCGCAACTTCCCGGTGGCGCCAGCAGGTGTGATCGTACCTACCCCAAG	185
8	Db	GlnAlaLeuAspThrPheGluAlaArgHisAspAspIleValLeuAlaSerTyrProLys	27
186	QY	TCCGGCCACCAAC-----TTGTGTCAGGAGGTGGTCTACTTGGTGGACGACGGC	233
28	Db	CysGlySerAsnTrpIleLeuHisIleValSerGluLeuIleTyrAlaValSerLys---	46
234	QY	GCTGACCCCGGATGAGATCGGCTTGATGAACATCGACGACGCTCCCGGTCTGTGAGTAC	293
47	Db	-----LysLysTyrLysTyr	51
294	QY	CCACAG---CCGGCCCTGGACATC-----ATCAGGAGCACTG	326
52	Db	ProGluPheProValLeuGluCysGlyAspSerGluLysTyrGluArgMetLysGlyPhe	71
327	QY	ACCTCTCCCCCCTCATCAAGAGCCACCTGCCTACCGCTTTCTGCCTCTGACCTCCAC	386
72	Db	ProSerProArgIleLeuAlaThrHisLeuHisTyrAspLysLeuProGlySerIlePhe	91
387	QY	AATCGAGACTCAAGGTCTATATATGGCTGCACCCCAAGGATCTGGTGTGCTCTAT	446
92	Db	GluAsnLysAlaLysIleLeuValIlePheArgAsnProLysAspThrAlaValSerPhe	111
447	QY	TATCAGTTTCCACCGCTCTCTCGGACCATGACTACCGAGGCACCTTTTCAAGAAATCTCGC	506
112	Db	LeuHisPheHisAsnAspValProAspIleProSerTyrGlySerTrpAspGluPhePhe	131
507	QY	CGGAGGTTTATGAATGATAAGCTGGGCTACGGCTCCTGG	545
132	Db	ArgInPheMetLysGlyGlnGluSer---GlyCysTrp	143

RESULT 71

US-09-252-991A-22452
; Sequence 22452, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:

```
Db 152 AlaProAlaProArgArgProAlaValProGlySerTrpArgCysArgAlaGly 171
Qy 135 -----CGATCTCTCCATCTTCC 118
Db 172 ProValArgAlaAlaProAlaProGlyArgGlyTrpArgArgSerProProAlaAla 191
Qy 117 -----CGCGGAGAGGGCGCGCGCA----- 94
Db 192 AlaThrAlaProAlaAlaGlyAlaGlyArgThrSerAlaGlyAlaAlaProAlaGlyArg 211
Qy 93 -----CGCCATGGAAGTCCGAAGTACTTGC 70
Db 212 ProAlaAlaProProAlaAlaAlaProGlyArgAsnArgProGlySerArgArgThrAla 231
Qy 69 TCTCGAAGTCCCGGGGTGCTGGGGTCTCGGCTCTCGCTCCGCGATGCGCGCGGT 10
Db 232 GlyArgProArgProGlyArgTrpArgThrArgHisArgGlyArgProAlaAlaAlaGly 251
Qy 9 CGCGGTGCG 1
Db 252 ArgAsnArg 254

RESULT 72
US-09-252-991A-27480
; Sequence 27480, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27480
; LENGTH: 313
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27480

Alignment Scores:
Pred. No.: 2,85e-08 Length: 313
Score: 176.00 Matches: 80
Percent Similarity: 31.4% Conservative: 11
Best Local Similarity: 27.6% Mismatches: 115
Query Match: 4.0% Indels: 84
DB: 2 Gaps: 11

US-10-768-158-1 (1-2419) x US-09-252-991A-27480 (1-313)
Qy 753 GGGCAGGGCTCAGCGTTGCAGCACTGGTCCACACGCTGT---GGCAGTGTCCGTCA 697
Db 32 ArgAlaGlyProArgArgSerProProGlySerProAlaSerArgGlyCysArgProArg 51
Qy 696 GGGCTTCAGCTGGGCTTGTTCACAGGACACCCCGAGGAATCTGGCCAGCTGTCTCCACCA 637
Db 52 ArgGlySerAlaGlyProAsnArgAsnProAlaProGlyArgArgProGlyTy:ProAla 71
Qy 636 ---TCGTACACAGGTCGGATGCATGCTTTCATCTTGAGAAAAGACGCTTCAGTCCA 580
Db 72 ArgSerAlaProGlyProProAla-----SerAlaAlaAlaSerAlaPro 86
Qy 579 TGCGGT---GCTCCAGAACTCTCGACGT-----GCTCAAAACGAGG 541
Db 87 ProGlyCysValProProAlaAlaSerArgAlaGlyAspTrpProAlaAlaArgAsnAsp 106
Qy 540 AGCGGTAGCCGAGCTTATCATTCATAACCTCCGCGAGAAATCTTTGAAGGTGCTCGGT 481
Db 107 SerArgArgPro-----HisArgGlyValProGlyArg-----ProGlyCysArgGly 122
```

```
Qy 480 AGCTCATGGTCCGACAGAGCGGTGGAACTGATATAAGACACCACGAGTCTTGGGT 421
Db 123 AlaProAlaSerAlaSerMetAspSerValSerAla----- 134
Qy 420 TGCAGCCATATAGATGACCTTGGAGTCTCCATTTGTGGAGTCCAGAG----- 373
Db 135 -----SerGlnHisAlaGlyGlnAlaArgProThrHisArg 146
Qy 372 -----GCAGAAAGCGTAGGCGAGTGGCTTCTTGATGAGCGGGAG 331
Db 147 GluAlaProArgAlaAlaProHisSerGlyArgAla----- 158
Qy 330 AGTCACTTCTTGATGATGTCAGGCCCGGCTGTGGTACTCCAGGACCGGAGTGTCT 271
Db 159 -----AlaAsnProGlyArgGlyAlaAla 166
Qy 270 CGTCGATGTTTCATCAAGCCGATCTCATCGGGTCCAGCGCTGGCTCACCAGTAGACCA 211
Db 167 ArgProAlaAlaArgProGlySerAlaProAlaAspProProAlaGluSerProAlaPro 186
Qy 210 CTTCTCGAAGACTGGTCCCGACTTGGGTAGGTGACGATCCACAGCTGCTGGGCC 151
Db 187 AlaProAlaSerGlyAlaAlaArgProAlaProArgProGlyProArgHisAlaAla 206
Qy 150 GCACCGGGAAGT----- 139
Db 207 ValAlaGlyAlaArgAspCysProAlaArgGluSerAlaValAlaAlaThrGlyValPro 226
Qy 138 -----TGGCCATCTCTCCATCTTCCCGGCGAGAGGGCGGCGCACGC 91
Db 227 AlaArgArgCysTrpArgArgProAlaProAlaGlyThrGlyAlaAlaProAlaPro 246
Qy 90 CATGGAACCTCGAAGTACTTCTCTCGAACTCCCGGGGTGCTGGGGTCTTGGCCCTCGC 31
Db 247 GlyCysProValProAlaGlyGlyArgProProCysTrpArgGlyArgProArg 266
Qy 30 TCTCCGCGCATCGCGCGCGCTGCCCTGCG 1
Db 267 ProGlyAlaHisArgArgProAlaArgLys 276

RESULT 73
US-09-252-991A-28620
; Sequence 28620, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28620
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28620

Alignment Scores:
Pred. No.: 3,26e-08 Length: 267
Score: 175.00 Matches: 86
Percent Similarity: 32.8% Conservative: 13
Best Local Similarity: 28.5% Mismatches: 111
Query Match: 4.0% Indels: 92
DB: 2 Gaps: 17

US-10-768-158-1 (1-2419) x US-09-252-991A-28620 (1-267)
```

Qy 783 CTTTCCACAGCCAACTCTTCCCGGCCCAACGGGAGGCGCTTCACGGTTGACGCACTGGT 724
Db |||||
Qy 15 ProSerThrAlaGly-----ProGlyHisProGlyHisThrAlaGlyCysAlaGlyHis 32
Db |||||
Qy 723 CCACAGCTGGTGGAGTCTCCGTCAGGCTTCCAGCTGGCCCTTGTACAGG---ACA 667
Db |||||
Qy 33 ArgGlyArgAlaGlyProAlaProGlyTrpSerPro-----HisArgAlaAla 48
Db |||||
Qy 666 CCCCAGGAATCTGGCCAGCTCTCCACCATCGTCACCAAGTCCCGATGTCATGT----- 613
Db |||||
Qy 49 SerProAlaGlyTrpProGlySerThrAlaArgSerProGlyProGlyAlaGlyCys 68
Db |||||
Qy 612 -----CTTCATCTTGAGAAAGAACGACGTCGAGTCCAGTCCGCGGTGCT 571
Db |||||
Qy 69 TrpProGlyArgThrArgHis-----ArgAlaAlaArgGlyThrProAlaAsnTrp 86
Db |||||
Qy 570 CCCAGAACTCTGACGCTCAACACAGGAGCGGTAGCCAGCTTATCATTAACACC 511
Db |||||
Qy 87 ArgArgProProAlaProGlyArg----- 94
Db |||||
Qy 510 TCCGCGCAGAAATCTTTAAAGGTGCTCGTGTAGCTCATGTCGCGAGAGCGGTGGAAC 451
Db |||||
Qy 95 SerSerArgValArgArgSerCys---SerProProThrSerAlaCysArgProAlaThr 113
Db |||||
Qy 450 GATATAAGACACCCACGATCCTTGG-----GGTTGCGAG--- 415
Db |||||
Qy 114 GlyArgHisArgProGlyArgProTrpArgArgArgAlaProProGlyCysAlaArg 133
Db |||||
Qy 414 -----CCATATAGATGACCT-----TGGAGTCTCCATGTGGAGTTCAG 376
Db |||||
Qy 134 AlaAlaProAlaGlyProSerArgTrpProArgArgGlySerArgProProGlyGlu 153
Db |||||
Qy 375 AGGCGAGAAAGCGGTAGGCGAGGTGCTTTCATGAGGCGGGAGAGTTCCTCTGA 316
Db |||||
Qy 154 ProAlaArgArgGlyCysAlaGlyAlaSer----- 163
Db |||||
Qy 315 TGATGTCAGCCCGGCTGTGGGTACTCCAGACCGGAGTGTCTCGATGTTTCATCA 256
Db |||||
Qy 164 -----ArgArgProArgArgThrAlaProGlySerGly----- 174
Db |||||
Qy 255 AGCCGATCTCATCGGGTCAGCGCTTGC-----TCACCAAGT 217
Db |||||
Qy 175 -----ArgProAlaArgProGlyArgThrAlaCysAlaArgAlaAlaAlaThr 190
Db |||||
Qy 216 AGACCACTCTCTGCA-----GCAAGTGTGTGCGGACTTGGGGT 178
Db |||||
Qy 191 ArgProAlaProAlaAlaProAlaGlyArgThrLeuAlaAlaCysArgProArgArg 210
Db |||||
Qy 177 AGGTGACATCCACA-----CGTCGTGGCCGCGCACCGGGAAGTTGGCACTCT 130
Db |||||
Qy 211 ArgArgArgThrThrGlyArgArgArgArgTrpSerAlaProAla-----ArgArg 228
Db |||||
Qy 129 CTTCCATCTTCCCGGCGAGAGGCGCGCAGCCGACCCATGGAAGTACTTGC 70
Db |||||
Qy 229 ProProGlyTrpArgSerGlyThrGlyArgProArgArgThrAlaProProAlaAlaCys 248
Db |||||
Qy 69 TCTCGAACTCCCGCGGTGTGGGGTCTCGGCTCTCTCCGCCATGCGCGCGGT 10
Db |||||
Qy 249 ArgArgAlaProArgGlyAla-----ArgArgAlaThrGlyCysArgTrpArg 264
Db |||||

RESULT 74
US-09-252-991A-24634
; Sequence 24634, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24634
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24634
Alignment Scores:
Pred. No.: 4,55e-08 Length: 320
Score: 174.00 Matches: 94
Percent Similarity: 31.0% Conservative: 23
Best local Similarity: 24.9% Mismatches: 125
Query Match: 4.0% Indels: 135
DB: 2 Gaps: 19
US-10-768-158-1 (1-2419) x US-09-252-991A-24634 (1-320)
Qy 1 CCGAGCGGCGCGCGCGCGCATGGCGGAGA-----GCAGGCGCGAGACCC 45
Db |||||
Qy 5 SerThrAlaProSerThrProCysArgArgProSerAlaSerAlaAlaGlyProArgThr 24
Db |||||
Qy 46 CCAGCACCCCGGGGAGTTCGAGAGCAAGTACTTCG-----AGTTCCATGGCGGCGCG 99
Db |||||
Qy 25 Pro---ProArgThrSerPheSerAlaSerArgProArgValThrSerProThrCysGly 43
Db |||||
Qy 100 TGCGCGCCCTTCTGCGCGGGAAGATGGAGGAGATGCCCAACTTCCCGGTGGCGCCAGCG 159
Db |||||
Qy 44 AspCysAlaArgSerThrSerArgTrpArgThrSer----- 55
Db |||||
Qy 160 ACGTGTGATCGTCACCTACCCCAAGTCCCGCACCGAGCTGTGTCGAGGAGTGTCTACT 219
Db |||||
Qy 56 -----SerProAlaProSerProCys----- 63
Db |||||
Qy 220 TGGTGAGCCAGCGGCTGACCCCGATGAGATCGGCTTGATGNACATCGACGAGCAGTCC 279
Db |||||
Qy 64 -----AlaSerAlaSerAla----- 68
Db |||||
Qy 280 CGGTCTCTGGAGTACCCACAGCGCGCTGGACATCATCAAGGAACGTGACCTCTCCCGCC 339
Db |||||
Qy 69 ArgSerGlySerSerArgAla-----Cys 78
Db |||||
Qy 340 TCATCAAGAGCCATCGCTCCCTACCGCTTTCTGCGCTCTGACCTCCACAAATGGAGACTCA 399
Db |||||
Qy 79 AlaSerThrAlaSerCysSerProIleSerProAlaCysArgAlaThrAlaProPro 98
Db |||||
Qy 400 AGGTCACTATATGGCTCGCAACCCCAAGGATCTGGTGTCTTATTTATCATGTTCCACC 459
Db |||||
Qy 99 ArgArgSerProGlyThrAlaProArg----- 108
Db |||||
Qy 460 GCTCTCTGCGGACCATGAGCTACCGAGGACCTTTCAAGAATTTCTCGCGAGGTTTATGA 519
Db |||||
Qy 109 ---AlaCysThrAlaArgAlaThrSerThrAlaIleArgArgSerAlaCysGlyMet--- 126
Db |||||
Qy 520 ATGATAAGCTGGGTACGGCTCTCGGTTTGTAGACGTCGACGAGTTCTGGGAGGACCGCA 579
Db |||||
Qy 127 -----ProTrpLysSerAlaMetArgAlaSerSerThrArgArgSerSerSer----- 142
Db |||||
Qy 580 TGGACTCGAAGCTGCTTTTCTCAAGTATGAACATGCATCGGAGAC-----TGG 630
Db |||||
Qy 143 -----CysCysAlaAsnThrGlyLeuProTrp 151
Db |||||
Qy 631 TGACGATGGTGGAGCAGCTGGCCAGATCTCTGGGGTGTCTCTGTGCAAGGCCAGCTGG 690
Db |||||
Qy 152 -----TrpTrpProIleProProAlaSerGlyArgCysSerAlaThrSer----- 166
Db |||||
Qy 691 AAGCCCTGACGAGCAGTGCACCGAGCTGGTGGGAC-----AGTGTGCAACGCTG 741
Db |||||

Db 167 AlaProThrSerSerThrCysAlaCysMetAlaThrArgSerSerThrAlaAlaAlaIle 186
Qy 742 AGGCCCTGCCGTCGGCGGGGAGAGATTGGCTGTGAAGACATCTTCAACGCTCTCCA 801
Db 187 ArgIysArgProTrpAlaSerGlyArgHisGly-----SerAlaProGlyPro 202
Qy 802 TGAATGAGAAGTTTGACTTGGTGATATAAC-----AGAAGATGGAAAGT 846
Db 203 AlaAlaThrArgProThrThrCysAlaGlyProAlaArgThrProAlaArg- 219
Qy 847 GTGACCTCACGTTTGACTTTTATTATAATAACAGAAACAACACCTGCATGCTCACAAAT 906
Db 220 -----ArgArgArgArg-AlaThrCysThrAlaThr-- 229
Qy 907 ACCCAGACAGTCTACTAGCCAAAAGTCCTGATGATCATTTATTCCT 955
Db 230 -----SerThrThrThrSerArgCysAlaHisProThrThrProGlyAlaSerCy 246
Qy 956 ---TGCTGGACAACTCTGGAAGCAGCGTGTGAACACAGCGGGGAAGAGCGCG 1011
Db 246 sSerCysTrpAenSerThrAlaAlaCysAla----- 256
Qy 1012 TGAGCGAGGAGTGTGATGATTCCTCAACCGAAGCAGCGTGTGCGCCTT 1060
Db 257 -----ArgCysProAspSerCysProMetThrCysGlyArgVal 269
RESULT 75
US-09-252-991A-28001
; Sequence 28001, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIORITY FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28001
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28001

Alignment Scores:
Pred. No.: 7,8e-08 Length: 657
Score: 173.50 Matches: 184
Percent Similarity: 27.1% Conservative: 44
Best Local Similarity: 21.9% Mismatches: 204
Query Match: 3.9% Indels: 410
DB: 2 Gaps: 42

US-10-768-158-1 (1-2419) x US-09-252-991A-28001 (1-657)

Qy 2256 AGAAGCAGCCTCCCTCAAGAAACACAGGCAAGGTGCTGCGAGCCCATCGATGC 2197
Db 121 ArgArgHisProProCysArgGluArgGlnAla----- 132
Qy 2196 CACGTGCCACCTGCTGCGCAGCTTCTCCCATTTACAAACAAACA----- 2149
Db 133 ---ValProGlyGlnProAlaAlaAlaGlnLeuGlnAlaArgAlaHisArgLeuPro 151
Qy 2148 -----CAAGCTTCTCTACACCCGCTTTCAATACAGCACCAAGAACGA 2104
Db 152 ArgProArgLeuAspArgAlaArgLeuArgGlnArg----- 163
Qy 2103 GACCCCTTCCACAGCAGGACCCCTCTGACCCACCACTTTGGGTGCTTCCAGCCACATAA 2044
Db 164 -----SerAlaPro----- 166

Qy 2043 ATGACACACAACCCCGTCCCAACGGGTTGGCTACTGGCCCCCGGACTCTCCATCG 1984
Db 167 ---AlaGlnArgProAspProAlaGly----- 175
Qy 1983 CTTCAGCAATGTGGTGCAGAAAATTAGGAGTGAAGCAATAAAGACCTCACCTTTACTC 1924
Db 176 ArgGlnArgAlaGlyValArgSerLeuCysAlaSerGlyLeuArgThr----- 191
Qy 1923 CCTCAGATACACGAGGAGTGTGGACAGACACGAGCGGCGCGTCCCAACCTGATCC 1864
Db 192 -----GlyAlaArgHisLeuAsp-Arg-----Pr 199
Qy 1863 GAGTGAACACGGGTACACTCGCAAAATGGTCTCCACCGCTGAGCGCTCAGGTGCAC 1804
Db 199 oglyGlnArgProArgArgSerHisProGlyPhePro----- 211
Qy 1803 AGAACAATCGAATGATAAAGACCCGGGATTTTAGAAAATTTACTATTAAAGCATCTGCTT 1744
Db 212 -----ArgArgArgAlaHisArgArgLeuLeuSerAla----- 222
Qy 1743 GCCAGCATTTAGTGGCTATCATCTTCAGAAACAAGTAAGGCTGTGGAATTCAGCTCAA 1684
Db 223 -GlnArgLeuValGly----- 227
Qy 1683 TTCCGAGCGCAACTCGGTCCTTTGGTGGCGGACTCTCCACCCACTG----- 1638
Db 228 ---AlaArgHisProGlyLeuGlyIleProAlaAlaArgProValProValGluGluLe 246
Qy 1637 -----CATGCAAAATGCTCTCTGACTGCTGACTGATGACTCC 1606
Db 246 uHisHisGlnArgPheAlaLeuGlyHisArgSerProGlyAlaValProProAr 266
Qy 1605 ACCTGCCAGCTCACAGGAGAGCTTCCTCAGCACATGCT----- 1565
Db 266 gProAlaSer-AlaProArgArgProAlaAlaAlaAlaValSerLeuArgGlnA 286
Qy 1564 -----CGCGGGCTCTCTGCAATGGGCTTTGGAGCGCGCTGGAGTCTCTGG 1513
Db 286 rgProGlyProArgArgLeuArgHisArgAlaGlyGlyProAlaHisArgGlyProA 306
Qy 1512 CAGAAGGACGTCCTCTGAAAGACCTAGTGAAACAATTTCTGCTCATTTTCAGGAAGTTCT 1453
Db 306 laArgGluArgPro----- 310
Qy 1452 GAAGAAGTGAACCTCGTCAGATAAACAAGTGTCCAGATTCCCCCAACACACCACTC 1393
Db 311 -----AlaAlaAlaProAspProGluGlnGlnProGluH 323
Qy 1392 ATCCTCTGTTCTCACCCTATCTCAGAACCAAGTGAGAACAGATGGCTGTGCTTAGAT 1333
Db 323 isValLeuAspGlyGlyAlaAspGlyArgProProGlnArgGlnTrpLeuProVal----- 341
Qy 1332 AATAAACTACAAAAATAAAGAGTTAACCTGAGATCTTTCTACCATTCGGGTGTGGCTC 1273
Db 342 -----AlaAlaA 344
Qy 1272 GCTCTGATTCCTTGGAAATGAACCTTTTATTTGTTTACTGACATTTATGTAGATTTC 1213
Db 344 rgArg-----ProValArgLeu-----ArgHisP 352
Qy 1212 CAGTGAAGGCTCTATAAATAACAATAACGCGGTTGAAAAGGCAGACATTTCTAGTTGC 1153
Db 352 roValArgArg----- 355
Qy 1152 ATATATTACAGGCTTTTCTTACGTCAGGCCATTGGAACCTGCAATGTGGAGACTGTT 1093
Db 356 -----ArgAlaArgProValArgLeuAlaAlaGlyAsp----- 366
Qy 1092 TGTAAATCAGCATGGAGA-----GGCTGCAGGTTCTAAAGCGGAGACAGCTGCTTC 1042
Db 367 ---His-ArgTrpArgGlnAlaAlaGlyAlaThrAlaSerArgArgAspProAlaPhe 384

```
Qy 1041 GGTGGGAATCATCATCCCTCCGCTCAGCCGCTCTTCCCTTCCCGCTG----- 988
Db 385 -----LeuGlyArgArgArgGlyHisProProAlaProAla 397
Qy 988 ----- 988
Db 398 AlaSerAlaArgAepMetPheArgSerValSerAlaAsnValAlaAlaLysCysSerArg 417
Qy 987 -----TTTCACACGCTGCTTCCA 970
Db 418 ProIleArgGlyArgAlaAspGluAlaLeuTyrlLeuLeuProPheHisLeuLeuPro 437
Qy 969 GAGTTTGCAGCAAGGAATAAATGAATGCATACAGACTTTTGGCTAGTACTGTCTG 910
Db 438 GlyAlaHisArgProGlyAlaGlu----- 445
Qy 909 GGTATTGTGAGCATGCAGGTTGTTCTTCTTATTATAATAAAGTCAAACTGTAGGT 850
Db 446 -----Gly 446
Qy 849 CACACTTTCCCATCTTCTGTTTATACACCAAGTCAAACTTCTTCTATTCATGAGACGGTGA 790
Db 447 AlaGlyLeuProValAlaAlaGlyGlnProGlyPro-----ArgArgArg 461
Qy 789 AGATGCTCTCCACAGCCCAACTCTTCCCGGCCCAACGG-----GCAGGCGCTCAGCGTTGC 733
Db 462 ArgThrProProThrArg-----LeuProGlyAlaGlnProAlaGlyProArg----- 477
Qy 732 AGCACTGTCCACAGCTGTGGCAGTCTCCGTCAGGCTTCCAGCTGGGCTTGTCTAC 673
Db 478 ---ThrGly-----ProAlaGlyArgArgGlyArgAlaAlaAspProValAlaGlyAspHis 495
Qy 672 AGGACACCCCGAGGAATCGGCCAGCTCTCCACCA-----TCGTCACCA 628
Db 496 ArgValProArgGlyAlaLeuProAlaProArgProAlaLeuGlnArgSerAlaAlaPro 515
Qy 627 GGTCCGATGATGTCTTCATCTTGAATAAGACGTTCCAGTCCAGTGGGCTGCC 568
Db 516 ArgPro-----ArgAlaArgSerArgGlyAlaGlyArgLeu 527
Qy 567 AGAAGTCTGACGCTCAACACGAGGCGTAGCCAGCTTATCATTCATAAACCTCC 508
Db 528 ArgHisProProValAlaGlnArgGlnArgAlaGlnProAla----- 541
Qy 507 GGCAGAAATCTTGAAGGTGCTCGGTAGCTCATGTCGCGCAGAGCGGTGGAAGTAT 448
Db 541 ----- 541
Qy 447 AATAAGACACCCAGATCTTGGGTTGCGAGCCATATAGATGACCTTGGAGTCTCOAT 388
Db 542 -----AlaProValGlyThrArg----- 547
Qy 387 TGTGAGGTACAGGGCAAGCGGTAGGCGAGGTGCTCTTGATAGGCGGGGAGAGG 328
Db 548 -----ArgArgAla----- 550
Qy 327 TCAGTCTCTGATGATGCCAGGCGCGCTGTGGTACTCCA-----GGACCGGAGCT 274
Db 551 -----GlyProAlaValAspArgProLeuGlyArgProGlySer 563
Qy 273 GCTGTCGATGTTTCATCAAGCCGATCTCATCGGGT-----CAGCGCCCTGGCTCA 223
Db 564 ArgArgArgAlaThrAspArg-----ArgProGlyLeuValLeuLeuArgProAlaArg 582
Qy 222 CCAAGTGA-----CCACCTCTCGCAGCAGCTGGTCCGAGCTTGGGTAGGTAGCA 169
Db 583 ProGlyArgArgLeuProGlyProAlaAlaLeu-----ArgArgArg 596
Qy 168 TCCACACGCTCGTGGCGCCGACCGGAAGTTGGCGA-----TCTCTCTCATCTTCC 118
Db 597 AlaLeuArgArg-----GlyAlaGlyArgLeuAlaThrAspProProGlyArg 612
Qy 117 CGC---GGCAGAGGGGGCGACCGCCATCGAACTCGAAGTACTTGTCTCGAACT 61
```

```
Db 613 ArgProGlyArgArgPro-----SerGlyVal 621
Qy 60 CCGCGGGGTCTGGGGTCTCGGCTCGCTCTCGCCATCGCGCGCGCTCGCGTGC 1
Db 622 ProProGlyProSerGly---GlnProAlaGlyHisProGlyArgLeuArgAspArg 640

RESULT 76
US-09-252-991A-31760
; Sequence 31760, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31760
; LENGTH: 1706
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31760

Alignment Scores:
Pred. No.: 1,72e-07 Length: 1706
Score: 172.50 Matches: 106
Percent Similarity: 35.3% Conservative: 26
Best Local Similarity: 28.3% Mismatches: 98
Query Match: 3.9% Indels: 144
Db: 2 Gaps: 24

US-10-768-158-1 (1-2419) x US-09-252-991A-31760 (1-1706)
Qy 762 CCGCGGCCA-----CGGCGAGGCTCTAGCGTTGCAGCACTGGTCCACCA 718
Db 223 ProAlaProAlaArgArgThrAspArgAlaGlyGluArgLeuAlaAlaGlyLeuPro 242
Qy 717 GCTGGTGGCAGTGTCCGCTCAGGCTTCCAGCTGGGCTTGTCTACAGACACCC----- 664
Db 243 -----GlyProAlaProAla-----ProGlyAlaGlyHisArgGlnProGlyAla 258
Qy 663 CCAGGA-----ATCTGGCCAGCTGCTCCACCATCGTCACAGGTCCCGATGCTT 610
Db 259 ProGlyGlyLysLeuArgAlaAlaGlyProProLeuAlaThrGlyPro----- 274
Qy 609 CATACTTGAGAAAAAGCACGTTCCAGTCCATGCGGTGCTCCC----- 568
Db 275 -----LeuArgArgGlnProAlaGlyAlaThrAlaArgLeuArgHisArg 289
Qy 567 -----AGAACTCTCTCAGCTGCTC----- 549
Db 290 ProGlyHisProAlaArgLeuArgArgGlnProAlaArg-LeuArgArgSerGlyGluAr 309
Qy 548 -----AAACACGAGGCGCT 535
Db 309 gProAlaGluSerProArgArgThrProGlyGluLeuArgGlyArgArgProAspAlaLe 329
Qy 534 AGCCGAGCTTATCATTAACCT-----CCGCGACAATTTTGAAGGTGCCTC 484
Db 329 uProArgLeu---AlaGlyLysProGlyArgTyrlProAlaGluPheAlaArgGlyAlaAl 348
Qy 483 GGTAGCTCATGTCGCGAGAGCGGTGGAAGTCACTGATAATAAGACACCAAGATCTTGG 424
Db 348 aProAlaProGlyProLeuAlaAlaAla-----ValProArgArgHis----- 362
Qy 423 GGTTCGAGCCATATAGATGACCTTGGAGTCTCCATGTTGTGAGGTTCAGAGGGCAGAA--- 367
Db 423 -----ValProArgArgHis----- 362
```

Db 363 ----AlaArgAlaAlaGlu--ProTyrArgArgProAlaGlyGlyArgAlaAspArg 380
Qy 366 ::::ACGGTAGGGCAGGTGCTCT 346
Db 381 ThrAlaThrAlaProAlaArgProArgThrAlaCysAlaAlaValAlaGlyProArg 400
Qy 345 TGATAGCGGGAGAGAGTTCCTTGATGATGATCCAGGCCG--GCT 298
Db 401 ----HisAlaArgA-gluGlnProArgThrLeuProGlyProArgProGlyAla 418
Qy 297 GTGGGTACT-----CCAGGACCGGGAGGTGCTCGTCGATGT 262
Db 419 ArgGlyThrHisLeuAlaGlySerAlaGlnProAlaAlaAlaArgProPro 438
Qy 261 TCATCA-----AGCCGATCTATCGGGTCAGGCCCTGGCTCACCAGTAGACCACT 208
Db 439 AlaAlaLeuProAlaArgAlaArgGlyArgArgGlnProProGlyArgProPro 458
Qy 207 -----CCTGCAGCAGC----- 196
Db 459 ArgArgLeuArgProGlyAlaProValAlaAlaAlaProAlaThrGlyGlyAlaAla 478
Qy 195 -----TGGTCCGGACTTGG-----GGT 178
Db 479 ThrGluProGlyAlaGlyAspSerArgHisArgCysArgSerGlyThrAspArgProGly 498
Qy 177 AGGTGACGA-----TCCACACGTCTGGCGCCGCGCAGCGGAAGTTGGCGATCTCTCCA 124
Db 499 ArgProArgGlnProAlaAlaAlaArgAspProGlyProGlySerTrpProGlyProSer 518
Qy 123 TCTTCCGCG-----GGCAGAAGCGCGCAGCCGACCAT 88
Db 519 AlaAlaArgArgAlaAlaArgArgGlyAspProHisArgArgGluThrAlaLeuArgSer 538
Qy 87 GGAACCTCGAAGTACTTGTCTCGAACT---CCCCCGGGGTGCTGGGGTCTCGGCCTCGC 31
Db 539 AlaAlaArgArgProArgArgAlaAspProProGlyArgThrProAlaGlnProGly 558
Qy 30 TCTCCGCCATGCCGCGCGCT-----CGCGGTGCG 1
Db 559 GlnThrGluGlyAlaArgArgLeuAlaArgArgArg 570

RESULT 77

US-09-252-991A-23674
; Sequence 23674, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23674
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23674

Alignment Scores:
Pred. No.: 9,25e-08 Length: 335
Score: 171.00 Matches: 89
Percent Similarity: 33.1% Conserved: 23
Best Local Similarity: 26.3% Mismatches: 110
Query Match: 3.9% Indels: 116
DB: 2 Gaps: 19

US-10-768-158-1 (1-2419) x US-09-252-991A-23674 (1-335)

Qy 840 CCATCTTCTGTTTATACACCAAGTCAAACTTCTCATTCATGAGACGGTGAAGATGCTCT 781
Db 15 ProSerSerProCysArgProAlaAlaThrAlaArgSer----- 27
Qy 780 TCCACAGCCCAACTCTTCCCGGCCACCGGCGGCTCAGCGTTCAGCACCTGGTCCA 721
Db 28 -----GlnGlyArgArgAlaGlyAlaProArgCysArgArgGlySer 41
Qy 720 CCAGCTGGTGGCAGTGTCTCCGTCCAGG-----CTTCCAGCTGGGCCCTGT--- 676
Db 42 AlaGlyAlaAlaGlyGlyCysArgGlyArgAlaArgArgSerProAlaSerProCysAla 61
Qy 675 -----CACAGG-----ACACCCCA-----GGATCTGG--- 652
Db 62 GlyArgHisArgArgAlaAlaValProAlaAlaArgArgArgTrpGlyThrTrpCys 81
Qy 651 -----CCAGCTGCTCCACCATGCTCACCAGGTCCC 622
Db 82 AlaGlySerArgGlyAlaAlaProGlyProValAlaArgProArgSerAlaProArgPro 101
Qy 621 GATGCATGTCTTCACTTCTGAGAAAACAGCTTCGAGTCCA---TCCGGTGTCTCCC--- 568
Db 102 AlaAlaArgSerGlyValProGlyArgArgCysArgProGlyCysProAlaProVal 121
Qy 567 -----AGAACTCTCTGACGTCTCAACCCAGGAGCGTAGCCAGCTTATCATTCATAA 514
Db 122 AlaArgGlyThrThrCysAlaAlaSerArgThrAlaAlaProAlaAlaArgGly--- 140
Qy 513 ACCTCCGCGCAGAAATCTTGAAGGTGCTCGGTAGC-----TCATGTCGCGCA 466
Db 141 -----ArgCysGlyGlySerArgArgGlyCysArgTyrValArg 153
Qy 465 GAGAGCGGT-----GGAACGTGATAATAAGACACACACAGATCTCTGG----- 424
Db 154 ArgProGlyLeuArgGlyProValArgArgArgProValArgTrpProArgThrAla 173
Qy 423 -----GGTTCCGAGCCATATAGATGACCTTGG 397
Db 174 ArgAlaThrAlaValArgArgArgProGlyGlyCys----- 186
Qy 396 AGTCTCCATTTGTGA-----GGTCAGAGGGCAGAAAGCGGTAGGCGAGGT 352
Db 187 AlaArgArgCysGlyTrpProAlaGlyAsnAlaArgAlaArgArgAsnArgAla 206
Qy 351 GGCTCTTGATGAGCGGGAGAGGTCTCTTGTGATGATGTCACGCGCCGCTGTGGGT 292
Db 207 GlyArg-----ProAlaAlaGly 212
Qy 291 ACTCCAGGACGGGAGCTGCTCTCGATGTTTCATCAAGCCGATCTCATCGGGTCAAGGC 232
Db 213 ArgAlaLeuArgSerArgSerArgArgCysAlaValAlaAlaArgSerAlaAlaArgArg 232
Qy 231 CCTGGCTCACCAGTAGACACCACTCTCTGACGACCAAGCTGGTCCGCGACTTGGGGTAGGTGA 172
Db 233 AlaArgAlaAlaArgProAlaProAlaAlaPro-----AlaTyrProAlaCys 249
Qy 171 CGATCCACACGTCTGGCGCGCACCGGAAAGTTGGCGATCTCTCT----- 127
Db 250 ArgSerGlyArgGlyAlaSerAlaProGlyAlaSerThrAlaProGlyProCysIlePro 269
Qy 126 -----CCATCTTCCCGCGCAGAAAGGGCGCGCAGCCGACCATGGAATCGA--- 79
Db 270 AlaValArgProValAlaArgArgAlaAspAlaGlySerAlaArgLeuProProArgArg 289
Qy 78 -----AGTACTTGTCTCTCGAACTCCCCCGGGGTCTGGGGGTCTCGGCCTCGC 31
Db 290 ArgGlnArgGlnCysSerThrArgProThrGlyCys-----ArgProArg 304

RESULT 78

US-09-252-991A-32035
; Sequence 32035, Application US/09252991A
; Patent No. 6551795

```
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32035
; LENGTH: 418
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32035

Alignment Scores:
Pred. No.: 1,32e-07 Length: 418
Score: 170.00 Matches: 82
Percent Similarity: 35.0% Conservative: 16
Best Local Similarity: 29.3% Mismatches: 118
Query Match: 3.9% Indels: 64
DB: 2 Gaps: 12

US-10-768-158-1 (1-2419) x US-09-252-991A-32035 (1-418)
Qy 798 AGACGGTGAAGATGCTCTCCACAGCCCACTCTCCCGCGCCAGCGGCGCTCAG 739
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 17 ArgArgProArgSerPro-----ProArgArgGlyAlaAla 28
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 738 CGTTGCAGCACTGGTGCACAGCTGGTGGCAGTGTCCGTCAGGGGCTTCCAGCTGGG--- 682
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 29 ArgAlaThrAlaArgProAlaLeuArgGlyProArgHnaAspArgProProAlaGlySer 48
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 681 -----CCTGTGCAGACAGACCCCGAGAACTCGGCAGCTGTCCACCATCG 634
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 49 SerSerSerProLeuSerArgArgValProAla-----AlaSerAlaGlyAla 65
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 633 TCACCAAGTCCC-----GATGCATGTCTTCATCTTGAGAAAGACAGCTTCAGATCCA 580
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 66 ThrProGlyProValArgAspGlyCysAlaGlySerValGlnArgProGlySerAlaArg 85
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 579 TCGCGTGCT-----CCGAGAACTCTCGCACGT 553
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 86 AlaGlyAlaSerSerProGlyArgSerArgValProGlyProProArgSerArgAspArg 105
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 552 GCTCAACACGAGGACCGTAGCCAGCTTATCATTCATAAACCTCGGC-----AGAATT 499
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 106 ArgArgValArgSerGlyArgGluSerArgHisGlyIleAlaGlyGlySerAlaArggly 125
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 498 CTTGAAAGGTGCCTCGGTAGCTCATGTGTCGCA-----GAGAGCGGTGGAACTGATAAT 445
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 126 TrpArgHisAlaGlyArgArgCysSerAlaProAlaThrGlyArgAspGly 145
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 444 AAGACACACACAGATCTCTGGGGTTGGCGACCATATAGATCACCTTGGAGTCTCCATTGT 385
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 146 SerArgProProHisProGlyAlaAlaAspArgProGlyAlaAlaAspArgArgCys 165
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 384 GGAGGTGAGAGGGCAGAAAGCGGTAGGCGAGGTGCTTGTATGAGCGGGGAGAGGTCA 325
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 166 ProGlyLeuProAlaSerProProArgArgAlaAla----- 178
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 324 GTTCTTGATGATGTCAGGCGCGGCTGTGGTACTCCAGGACGGGAGCTGCTCGTCA 265
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 179 -----ProArgProAlaThrGlyValAlaGlyProAlaAlaValAspArg 193
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 264 TGTTCATCAACCGCATCTCATPCGGGGTCAGCGGCTGGCTCA-----CCAAGTAGACCA 211
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 194 AlaAlaAlaArgGlySerCysArg---ProAlaProGlySerSerAlaProGlyValPro 212
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 210 CCTCCTGACGAAGCTGGTGCCGACTTGGGGTAGGTGACGATCCACACGCTGCTGGGCC 151
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Db 213 AlaProAlaAla-----ArgArgLeuProCysSerAla 223
Qy 150 GCACCGGGAAGTGGCGATCTCTCCATCTTCCCGCGGCAGA----- 109
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 224 AlaThrThrValGlyCysGlyCysProSerGlyThrGlyArgTyrArgHisAlaProPro 243
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 108 ---AGGGCGGCGACCGCACGCGCATGGAACTCGAAGTACTTCTTCGAACCTCCCGGGG 52
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 244 GlyProAlaSerAlaAlaArgAlaProAlaArgAlaAlaSerGlyThrProProGly 263
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
RESULT 79
US-09-252-991A-27249
; Sequence 27249, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27249
; LENGTH: 308
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27249

Alignment Scores:
Pred. No.: 1,24e-07 Length: 308
Score: 169.50 Matches: 80
Percent Similarity: 33.3% Conservative: 11
Best Local Similarity: 29.3% Mismatches: 95
Query Match: 3.9% Indels: 87
DB: 2 Gaps: 15

US-10-768-158-1 (1-2419) x US-09-252-991A-27249 (1-308)
Qy 762 CCGGCGCCAGCGGCGGCGCTCAGCTTGCAGCACTGGTCCACCACTGGTGGC----- 709
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 18 ProAlaArgGlyAlaGlyProGlyArgGlyGlySerGlyProGlyArgAlaGlyArgArg 37
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 708 AGTCTCGGTCCAGGGCTTCAGCTGGGCTTGTACAGGACACCC----- 664
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 38 SerIleProProAlaProAlaProArgArgGlyProArgLeuAlaArgProSerArgGlyArggly 57
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 663 -----CCAGGAATCTGGCCAGCTGTCTCCACCACTCGTCA 631
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 58 ArgThrArgAlaProGlyArgCysCysSerGlyTrpTrpSerThrGlySerProMetAla 77
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 630 CCAGGTCGCGATGCATGTCTTCATPACTTGAGAAAAGCACGTTCCAGTCCATCGGTGCT 571
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 78 ProValArgSerAlaCysArgGlyThrSerArgProValArgSerTrpPro---GlyPro 96
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 570 CCGAGAACTCTGCAGCTGTCAAAACCCAGGAGCGTAGCCAGCTTATCATTCATAAAC 511
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 97 ProAlaGlyArgAlaGlyGlyTyrGlyArgArgArgGlyPro----- 110
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 510 TCCGGCAGAAATCTTGAAGAGTGCTCGGTAGCTGTGTCGCGAGAGCGGTGGAAC 451
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 111 -----ArgArgSerValProProThrValAlaSerValAlaGly----- 123
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 450 GATAATAAGACACACAGATGCTTGGGGTTCGAGCCATATAGATGACCTTGGAGTCTC 391
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 124 -----ProSerArgSerGlyAlaGlyProAlaArg----- 133
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 390 CATTTGTGAGGTCTAGAGGGCAGAAAGCGGTAGGCGAGGTGGCTCTTGTATGAGCGGGGAG 331
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Db 134 -----ArgAla-----AlaGlyPro 138
Qy 330 AGGTCACTTCTTGATGATGTCAGGCCGCGCTGTGGGTACTCCAGGACCGGAGCTGCT 271
Db 139 GlyAlaAlaPro-----CysProGlyProGlyArgGlySer-----GlyAlaAla 153
Qy 270 CGTCGATGTTTCATCAAGCCGATCTCATCGGGTCAGCGCCTCGCTGCTCACCAGTAGACCA 211
Db 154 Arg-----ArgHisArgArgArgProGlyArgArgSerArg--- 167
Qy 210 CCTCTGCAGCAAGCTGGTGCAGACTTGGGGTAGTGACGATCCACACGTCGCTGGGCC 151
Db 168 -----TrpCysProCysAlaGlySer---ArgAlaArgProSerGly 181
Qy 150 GCACCGGGAAGTTGGCGATCTCTCCATCTTCCCGCGGCAGAAAGGGCGAGCGCAGCG 91
Db 182 CysThrGlyCysArgProArgLeuProArgAlaGlyArgArgAlaAla----- 198
Qy 90 CATGGAAGTCAAGTACTTGTCTCGAACTCCCGCGGGTGTCTGGGGTCTCGGCCCTCGC 31
Db 199 ---GlyGlyArgProProCys-----CysProGlyAlaProGlyGluArgLysArg 215
Qy 30 TCTCCG-----CCATGC 19
Db 216 SerThrGluSerSerAlaArgProAlaProArgProCys 228

RESULT 80
US-09-252-991A-24009
; Sequence 24009, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Matc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24009
; LENGTH: 697
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24009

Alignment Scores:
Pred. No.: 3,16e-07 Length: 697
Score: 167.50 Matches: 188
Percent Similarity: 30.2% Conservative: 53
Best Local Similarity: 23.6% Mismatches: 264
Query Match: 3.8% Indels: 293
DB: 2 Gaps: 43

US-10-768-158-1 (1-2419) x US-09-252-991A-24009 (1-697)
Qy 2059 CTTTCCAGCCACATAAATGACAGACAACCC-----CGGTCCCGCCGCGGTGGCTGGCT 2006
Db 2 MetSerSerHisAspProHisArgThrAlaLeuHisArgHisProArgProGlyThrThr 21
Qy 2005 ACTGGCCCCGACCTCTCCATCGCTTCACGGAATGTGGGTGAGAAAATTAGCGAGTGAAGC 1946
Db 22 LeuArgProArg----- 25
Qy 1945 AATAAGACCTCACCTTTACTCCCTCAGATACA-----CGGAGGATGAGGTGGACAGA 1892
Db 26 ArgArgAlaLeuProArgGlnProThrAspProLeuGlyArgArgGluGluAlaArgGly 45
Qy 1891 CACGAGCGCGCGTGCACCACTGATCCGAGTGAACAGGGGTACACTCGCAAAATGGTC 1832
Db 46 ArgAlaGlyGlyAlaAsp-----LeuArgAlaGlnGln-----GluArgArg 59
```

```
Qy 1831 CTCCACGGCTGAGCGCTCAGGTGCACAGAACAACTGAATGATAAAGACCCCGGGATT 1772
Db 60 AlaPro-----AspAlaSerGlyArgHis-----ArgArgProGly--- 72
Qy 1771 TTAGAAATTTACTATTAAAGCATCTGCTTGCAGCATTAGTGGCTATCATCTTCAGGAAC 1712
Db 73 -----AlaGluGlyProArg-----ThrGlyProGly 81
Qy 1711 AAAGTAAGGTGTGGAATTCAGCTCAAAATTCGAGCGCAACTCGTCTCTTTGGTGGCGGA 1652
Db 82 HisProArgAlaGlyProGlyArgGlnGluPro---ValGlyArgAlaAlaGluGlyArg 100
Qy 1651 CTCTCCACCCACTGCATGCAAAATGCTCTGACTGCATTAGACTCCCACTGCCAGCTCA 1592
Db 101 ArgHisLeuHisHisArgSerLeuProValProAlaProAspSer-----ProAlaAla 118
Qy 1591 CCAGGCAGACTTCCCTCAGCACATGCTCCGCGGGCTCCTCTGCAT---TGGGCCCTTTGGA 1535
Db 119 ProGlyArgAlaAlaAspAlaAlaValHisArgGlyGluLeuHisProTyr-ProProAr 138
Qy 1534 GCGCCCGCTGGAAAGTCTCTGGCAGAAAGAGCGTCCATTCTGAAGACCTAGTGAACAATTT 1475
Db 138 GlnAlaAlaTyrArgArgThrArgArgAspHisHisArgPro----- 152
Qy 1474 CTGCTCATTTTCAGGAAGGTTCTGAAGAACTCGTCCGTCAGATAAACAAGTCTCCAGA 1415
Db 153 -----AlaValProGlySerArgArgProAs 161
Qy 1414 TTCCCCCAACAGACACCACCACTCATCTCTGTCT--- 1381
Db 161 pGlnAlaAlaValArgArgAlaLeuLeuCysAlaAspAlaGlyArgProSerLeuAspAr 181
Qy 1380 -----CACACATCTCTCAGAACCAAGTGAGAACAG- 1351
Db 181 gGlnGlyGlnHisArgPheArgAlaAlaGlnArgGlnGluProAlaProAlaArgArgAr 201
Qy 1350 -----AATGGCTGTGCTTAGATAATAA 1328
Db 201 gProLeuLeuProArgProGlyAlaGlySerLeuAlaArgProCysAlaAlaThrAr 221
Qy 1327 AACTACAAAAATAAAGAGTTAAACCTCAGATCTTTCTACCATTCGGGTGTGGCTCGCTCC 1268
Db 221 gThrSerThrProArgTrpSerProArgArgTrpArgProSerAlaThrTrpLeuProPr 241
Qy 1267 TGATTCCCTTGGAAATGAACCTTTTATTTGGTTTACTGACATTTATGTAGATTTCCAGTG 1208
Db 241 oAsp-----TrpAlaSer--ArgCysCysArgSerPro--- 251
Qy 1207 AAAAGCTCTATAAATAACAATAATACGGGGTTGAAAAGGCAGACATTCTAGTTGCATATA 1148
Db 252 -----ArgTrpThrAlaThrThrArgArgAlaSer----- 262
Qy 1147 TTACAGCGTTTATCTTACCGTCCAGGCCATTGGAACTTGCAACTGGAGACTGTTGTGAA 1088
Db 263 -----SerLysCys----- 265
Qy 1087 TCAGACATGGAGAGGCTGCAGTTCTAAAGCGGAGACAGCTGCTTCGGTTGGGAATCATC 1028
Db 266 -----AlaArgSerAlaArgArgCysArgSerVal 275
Qy 1027 ACACCTCCCTCCGCTCAGCGCGCTCTCCCTTCCCGCTGTTTTCACACGCTGCTTCCAGA 968
Db 276 ProLeuPro-----SerProGlyAlaProAlaSerArgVal---HisAlaProSerLys 292
Qy 967 GTTTGTCCAGCAAGGAATAAATGAATGCATACAGGACTTTTGGCTAGTAGACTGTCTGGG 908
Db 293 CysTrpProThrArgSer----- 298
Qy 907 TATTGTGAGCATGAGGTTGTTGTTTCTGTTATTATTATAAATAAAGTCAACAGCTGAGGTCA 848
Db 299 -----AlaCys-----Ala 301
```

```
QY 847 CACTTTCCCATCTTCTGTTTATACACCAAGTCAAACCTTCTCATTCATCGAGACGGTGAAG 788
Db :|||||:
302 ArgTrpProVal-ArgLysProArgAsnArgLysSerHis----- 315
QY 787 ATGCTCTCCACAGCCCAACTCTTCCCGGCCACGGCAGGCGCTCAGCGTTGCAGCAC 728
Db :|||||:
316 -----AspArgAlaValGlnGlyProGlyTyArgAlaLysGlyArg----- 330
QY 727 TGGTCCACAGCTGTGGCAGTGCTCCGTAGGCTTCCAGCTGGCCCTTGT- --CACAG 671
Db :|||||:
331 ---ArgArgAlaGlyGlyGluAlaArgSerGlyArgProGlyAsnProAlaGlyTyPr 349
QY 670 GACACCCCGAGGATCTGCCAGCTGCTCCACCATCGTACACAGTCC- --GATGCATG 614
Db :|||||:
349 oValProProAla-----ProAlaLeuProGlyProTyPrProArgHi 363
QY 613 TCTTCATCTTGAGAAAAAGCACGTTTCG----- 586
Db :|||||:
343 sArgH:sArgArgProAlaAlaArgGlyArgArgGlyArgGlyArgGlyArgAr 383
QY 585 ---AGTCCATGCGGTGCTCCAGAACTCTCTGACGTGCTCAAACC---AGGAGCCGTAG 533
Db :|||||:
383 gArgArgGlyAspGlyAlaProProGlnProAlaGlyAlaProAlaGlyArgGlnArgHi 403
QY 532 CCCAGCTTATCATTCATAAACCTCCGGCAGAACTCTTGAAGGTGCCCTCGGTAGCTCATG 473
Db :|||||:
403 sPro-----GluProAlaLeuLeuProPheGl 412
QY 472 GTCGCGCAGAGCGGTGAAGTGAATAAGACACCA-----GATCC 428
Db :|||||:
412 nProGlyAlaGluGlyArgSerGlnAlaArgHisProProAlaLeuLeuArgArgSerPr 432
QY 427 TTGGGGTTGCAGGCCATATAGATGACCTTGGAGTCTCCATTGTGGAGTGCAGAGGCAGA 368
Db :|||||:
432 oAlaGlyCysLeuGly-----ProGlyAspLeuProSerGlyIleProArgAlaGl 449
QY 367 AAGCGGTAGGCGAGGTGCTCTTGATGAGGCGGGAG----- 331
Db :|||||:
449 uArgArgArgAlaAspSer-----GlyGlyAlaAspProAspAlaAspLeuProAs 466
QY 330 ---AGGTCAGTCTCTTGATGATGTCACAGGCCGCGCTGGGTACTCCA----- 286
Db :|||||:
466 pTyArgGly-----ProHisProAlaAlaProAlaProValGluProTh 481
QY 285 -GGACCGGGAGCTGCTCGATGTTTCATCAAGCGCATCTCATCGGGT----- 238
Db :|||||:
481 rGlyProGlyHisAlaArgProGlnGlnProAlaArgLeuAlaAlaGlyThrGlyPr 501
QY 237 -----CAGCGCCCTGGCTACCAAGTAGACACCTCTCTGAGCAAG 197
Db :|||||:
501 oArgLeuProThrTrpProAlaArgProGlyAspPro---LeuProAlaProAlaAla 520
QY 196 CTGGTGGCGGACTTGGGTAGGTGACGATCCACAGCTCGTGGGCC---GCACGGGNAAG 140
Db :|||||:
520 aGlyCysArgHisArgGlyThrGlyArgArgProAlaLeuGlyAlaValAlaProGlyLe 540
QY 139 TTGGCGA-----TCTCTCATCTTCCCGCGCGCAGGAAG 107
Db :|||||:
540 uArgArgThrAlaAspProProValIleProAlaAlaProAlaArgSerArgAlaLeuAl 560
QY 106 GCGCGCAGCCCGCAGCCATGGAATCGAAGTACTTGTCTCGAATCTCC----- 58
Db :|||||:
560 aGlySerAlaAlaLeuAlaGlyGlnProLeuAlaGluThrLeuProCysGlnProAr 580
QY 57 -----CCGGGTGTGGGGTCTGGGCTCGC----- 31
Db :|||||:
580 gLeuProAlaAspArgArgAlaThrProGlyArgArgGlyAspArgLeuArgProGlyPr 600
QY 30 -----TCTCGCCATGCGCGCGCTCGCGCTCGC 1
Db :|||||:
600 oGlyArgAlaAspAlaAlaProGlyAlaArgArgArgCysArg 615
```

RESULT 81

```
US-09-252-991A-30343
; Sequence 30343, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30343
; LENGTH: 286
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30343
```

Alignment Scores:

Pred. No.:	2,34e-07	Length:	286
Score:	166.50	Matches:	81
Percent Similarity:	33.9%	Conservative:	17
Best Local Similarity:	28.0%	Mismatches:	104
Query Match:	3.8%	Indels:	87
DB:	2	Gaps:	16

US-10-768-158-1 (1-2419) x US-09-252-991A-30343 (1-286)

```
QY 789 AGATGTCTCTCCACAGCCCAACTCTTCCCGGCCACGGCAGGCGCTCAGCGTTGCAGC 730
Db :|||||:
40 ArgCysArgArgGlySerAlaValAlaGlyCysArgAlaAlaSerArgArgSer 59
QY 729 ACTGGTCCACAGCTGTGGCGAGTGTCCGTCCAGGGCTTCCAGCTGGCGCTTGTCCAGG 670
Db :|||||:
60 IleSerProThrProGlyAlaAlaAla-----ArgArg 70
QY 669 ACACCCCGAGGAATCTGGCAGCTGCTCCACCATCGTCACCGAGTCCGATGTCATGCTT 610
Db :|||||:
71 SerAlaGlyGlySerTrpLysAla-----ProGlyArgCys--- 82
QY 609 CATACTTTGAGAAAAAGCACAGTTCGAGTCCATCGCT---GCTCCAGAACTCTCCAGCT 553
Db :|||||:
83 -----ProCysGlySerAlaGlyIleProProSerArg 93
QY 552 GCTCAAACAGGAGCGGTAGCCAGCTTATCATTCATATAACCTCCGGCAGAAATTTTGA 493
Db :|||||:
94 PheProAspAlaArgArgSerProGlyAlaAlaArgTrpProGlySerVal----- 111
QY 492 AGGTGCCTCGGTAGCTCATGTGTCGCGAGAGCGGTGGAATGATAATAAGACACCA 433
Db :|||||:
112 AlaAlaAlaAlaThrProTrpProArgGlyProArgProArgArgHisTrpArgProPro 131
QY 432 GATCCTTTGGGTTCCGAG-----CCATATAGA---TGACCTTGG 397
Db :|||||:
132 GlyProCysAlaAlaLysAlaAlaArgGlySerProAlaProTyArgTrpProProGly 151
QY 396 AGTCTCATTTGTGGAGTGCAGGGCAGAAAGCGTAGGGCAGGT-----GGC 349
Db :|||||:
152 SerValAlaAlaThrSerAlaArgAlaProProGlyArgArgAlaProGlyArgProGly 171
QY 348 TCTTGATGAGCGGGAGAGGTCAAGTTCCTTGATGATGTCAGGCGCGCTGTGGGTACT 289
Db :|||||:
172 SerArgThrGlyCysAlaArg-----ProAlaGlyCysProSerProAlaThrThrAla 189
QY 288 CCAGACCGCGGAGCTGCT-----CGTCGATGTTTCATCAAGCGCATCT 247
Db :|||||:
190 ProGlyProAlaAlaThrValProAspTrpArgAspProAlaArgCys---GlyArgArgSer 208
QY 246 CATCGGGGTGAGCGCCCTCGGTCA-----CCA 220
```



```
QY 733 CAGACTGTCACACGAGTGTGGCAGTGTCTCCGTCAGGGCTTCAGGTGGCCCTTGTC 674
Db 5 ArgArgTTPAlaSerSer-GlySerAlaSerAlaAlaArgArgProGlyArgLeuAr 24
QY 673 CAGGACACCCAGGAATCTGCCAGCTGCTCCACCATCGTCACACAGTCCCGATGCATG 614
Db 24 gArg-----ProGlyProProAspArgProGly----- 34
QY 613 TCTTCATCTTGAGAAAAGCAGTTCGAGTCCATCGCGTGTCTCCAGAACCTCTGCACG 554
Db 35 -----LysSerArgArgValValCys----- 41
QY 553 TGCTCAACACGAGCCGTAGCCAGCTTATCATTAACCTCCGCGCAGAACTCTTGA 494
Db 42 -----ProGlyTyrAlaArgArgThrSerAlaArg----- 51
QY 493 AAGGTGCTCGTAGTCTCATGTCGCCAGAGCGGTGGAACCTGATAATAAGACACACC 434
Db 52 ----- 52
QY 433 AGATCTTTGGGTTCCGAGCCATATAGATGACCTTGGAGTCTCCATTGTGGAGGTACAG 374
Db 52 oAspProGlyGlyCysSerPro---ArgGlyProArgProAlaArgArgThrGlyArgSe 71
QY 373 GGCAGAAAGCGGTAGGCGAGGTGC-----TCTTGATGAGCGCGGGAGAGGTACGT 323
Db 71 rAlaThrSerAlaSerSerAsnGlyArgCysProAlaArgLysAlaGlyProHisSerAl 91
QY 322 TCTTTGATGATCTCCAGGCCGCGTGTGGTACT-----CCAGGACCGGAGCTGCTCG 269
Db 91 aArgGlyAsnSerProAlaProAlaAlaGlyArgAspTrpProArgProGlySerGlyAr 111
QY 268 TCGATGT---TCATCAAGCCGATCTCATCGGGTCAGCGCCCTGCTCCACCAAGTAGACC 212
Db 111 gGlyCysProGlySerThrAspArgHisAlaGlyAlaArgProGly-----CysArgPr 129
QY 211 ACCTCTCCAGACAGCTGGTCGCGAGCTTGGGTAGGTAGATCCACACATCGCTGGGCG 152
Db 129 oAlaProSerAlaArgCysGlyArgArgTyrArgArgGlyLeuSer-----GlySe 146
QY 151 CGCACCGGGA---AGTTGGCGATCTCTCCCA-----TCTTCCGCGCGCAGAGG-- 106
Db 146 rGlyProGlyArgGlyThrArgArgProProHisArgAlaSerSerArgProArgArgSe 166
QY 105 -----GCGCGACCGCAGCCGATGGAATCGA---AGTAC 74
Db 166 rLeuAlaHisCysProValArgAlaAlaProAlaSerArgArgGlySerArgProAlaAs 186
QY 73 TTGC-----TCTCGAACTCCCGGGGTGCTGGGGGTCTCGGCTCGCTCT----- 28
Db 186 pCysArgTyrGlyArgThrProArgProAlaTrpArgArgAlaProGlyAlaSerSerAr 206
QY 27 -----CCGCCATGCCGCGCGT 10
Db 206 gArgAlaAlaAlaGlyCysProProProThrProArgArg 219
```

RESULT 84

```
US-09-252-991A-26942
; Sequence 26942, Application US/09252991A
; Patent No., 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26942
```

```
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26942
Alignment Scores:
Pred. No.: 3,37e-07 Length: 437
Score: 166.00 Matches: 95
Percent Similarity: 35.0% Conservative: 112
Best Local Similarity: 29.7% Mismatches: 117
Query Match: 3.8% Indels: 96
DB: 2 Gaps: 20
```

US-10-768-158-1 (1-2419) x US-09-252-991A-26942 (1-437)

```
QY 789 AGATGTCCTTCCACAGCCCAACTCTTCCCGCGGCCACGGCAGCGG-----CCTCAG 739
Db 6 ArgProProProValAlaGlnLeu---ProAlaGlyArgThrGlyLysProValProArg 24
QY 738 CGTTGCAGCACTGTGCCACAGCTGTGTGGCAGTCTCCGTCAAGGCTTCCAGTGGGCT 679
Db 25 ArgLysArgThrGlyLysProGlyAlaGlyAlaGluAspProGluGlnArgArgLysPro 44
QY 678 TGTCAAGGACACCCAGGAACTCTGCCAGTCTCCACCATCGTCACAGGTCCCGAT 619
Db 45 AlaArgArgLeuProProSerArgGlyAsnValGlnArg----- 57
QY 618 GCATGTCCTTATCTTCCAGAAAAGACAGTTCGAGTCCATCGCGTCT-----CCC 568
Db 58 -----GluArgCysArgGlnArgProPheArgAlaGlnCysGlnPro 71
QY 567 AGAACTCTGTCAGCTCAAAACAGGAGCGGTAGCCAGCTTATCATTTATAACCTCC 508
Db 72 GlnProProValLeuValGln-----ArgProAlaArgArgSerArgSer 87
QY 507 GGCAGA---ATTCTTGAAGGTGCTCGTAGCTCAGTCCGCGAGAGAGCGGT---GGA 454
Db 88 GlyGlnSerValValGlnArg-----AlaAspProGlyArgGly 100
QY 453 ACTCAATAAGACACACAGCATCTTGGGGTTGCGAGCCATATAGATGACCTTGGAGT 394
Db 101 GlyAspArgThr-----GluProLeuArgArgPro----- 111
QY 393 CTCCATTGTGGAGTCAGAGGCGAGAAAGCGGTAGGCGAGGTGGCTTTCATGAGCGGG 334
Db 112 LeuArgCysArgSerGlnSerProAlaAlaGlyArgGlyAla-----ProGly 127
QY 333 GAGAGTCACTTCTCT---TGATGATGTCAGGCGCGCTGTGGGTACTCCAGGACCGGGA 277
Db 128 ArgHisLeuLeuProGlyAlaGlnAlaProGlyProAlaArgArgThrGlyArgProAla 147
QY 276 GCTGCT-----CGTCGATGTTTCATCAAGCCGATCTCATCGG-----GGTCAGCGCCT 229
Db 148 AlaAlaProValArgArgAlaGlyGlyAlaGluArgArgArgSerGlyArgThrPro 167
QY 228 GGC-----TCACCAAGTAGACCA-----CCT 208
Db 168 GlyArgArgAlaGlyArgLeuArgProAlaLeuProGlyLysSerArgArgThrGluPro 187
QY 207 CTGTCAGCAAGCTGGTCCCGACTTGGGGTAGGTGACGATCCACACCTCGCTGGCGCGCA 148
Db 188 ProAlaProGlyArgGlyArgThrSerCysArgGlnArgGlyAlaGlyAsn---AlaAla 206
QY 147 CCGGGAAGTTGGCGATCT-----CCTCATCTTCCCGCGCGCAGAGGCG 103
Db 207 ProGlyHisAlaArgGlyAlaLeuArgHisArgProAlaSerArgArgGlnArgArgAla 226
QY 102 GCAGCCGCGCCCATGGA-----ACTCGAAGTACTTGC 70
Db 227 AlaGlyGlnArgThrGlyLysArgGlyValProGlyGlnArgGlnProArgAlaAlaAla 246
QY 69 TCTCGAACTCCCGGGGTCTGGGGTCTCGGCTCGCTCT----- 28
Db 69 TCTCGAACTCCCGGGGTCTGGGGTCTCGGCTCGCTCT----- 28
```

```
Db      247 GluArgAlaGlyGlnGlyArgLeuGlyArgArgThrValAlaAspGlnProGlyAspPro 266
Qy      27 -----CCGCATGCCCCCGCGCGTCGCGCGCGC 1
Db      267 GlyAspHisGlyThrAspLeuAlaHisProHisProGlyLeuArgArgSerArgArgArg 286

RESULT 85
US-09-252-991A-25699
; Sequence 25699, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25699
; LENGTH: 306
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25699

Alignment Scores:
Pred. No.: 3,06e-07 Length: 306
Score: 165.50 Matches: 87
Percent Similarity: 36.9% Conservatives: 20
Best Local Similarity: 30.0% Mismatches: 109
Query Match: 3.8% Indels: 74
DB: 2 Gaps: 18

US-10-768-158-1 (1-2419) x US-09-252-991A-25699 (1-306)
Qy      790 AAGATGTCCTCCACAGCCCACTTCTCCCGGCCACG-----GGCAGGGCC 743
Db      62 ArgArgSerArgGProProArgPheProArgProAlaArgTtpProGlyArgAla 81
Qy      742 TCAGGTTGCAGCACTGTCACACAGCTGGTGGAGTCTCCGTCAGCGCTTCAGCTGG 683
Db      82 ProAla---ArgHis-AlaProAlaAlaLeuArgArgAlaGlyGlySerProGlyGl 100
Qy      682 GCCTTGTCACAGGACACCCCGAGGA---ATCTGGCCAGCTGCTCCACCATCGTCACCCAGG 626
Db      100 YArgSerHisGlyTtpSerValGlyArgArgTtpProCysArgArgArgSerArgProAr 120
Qy      625 TCCCGATGCATGTCCTTCATCTTGAGAAAAGCAGCTTCGAGTCATCGCGTGTCTCCGAG 566
Db      120 gAlaGly-----ArgAlaArgGlyArgAenAlaGlyCysProAr 133
Qy      565 AACTCCTGCATGCTCAAAACGAGGCGGTAGCCAGCTTATCATTCATAAACCCTCCGG 506
Db      133 GArg---ArgArgGlyArgSerArgSerProArgGlyGlySerValProAlaGl 152
Qy      505 CAGATTCTGAAAGGTGCTCGGTAGCTCATGTCGCGCA-----GAGAGCGGTGGAAC 452
Db      152 YArg-----ArgTtpSerGlyAlaArgArgArgProGlyTh 164
Qy      451 T-----GATAATAGACACACCAACAGATCCTTGGGGTTGCGAGCCATATAGACCTTG 398
Db      164 rArgProArgProArGArgProProLeuProGlyArgThrAlaProAenArgCysArgTr 184
Qy      397 GAGTCTC-----CATGTGAGGTGAGAGGGCAGGAGCGG 362
Db      184 pCysValAspArgSerArgThrSerArgArgPheProAlaArgArgCysArgAlaGl 204
Qy      361 TAGGCGAGGTGGCTCTTGATGAGGCGGAGAGGTGACGTCTTCATGATGTCCAGGCCC 302
Db      204 YArgArgProValAla-----GlyGlyProArgArgProPro-----ProGlySe 219
```

```
Qy      301 GGCTGTGGGTACTCCAGGACCGGAGCTGCTCGATGTTTCATCAGCCGATCTCATCG 242
Db      219 rAla-----ProAlaAlaArgArgValArgCysArgThrCysAla----- 234
Qy      241 GGGTACAGCGCCTGGGTCCACCAAGTAGACACACCTCTCGACAGCAAGCTGGTGGCGGACTTG 182
Db      235 -----ArgProGlyGlySerArgArgArgAlaGlyGlyTyTfGlyAlaArgTr 252
Qy      181 G-----GGTAGGTGACGATCCACATCCACGTCGTCGTCGGCGCCGACC 146
Db      252 pArgTyrAlaCysArgSerArgProGlyArgCysArgAsnAlaArgArg-----Pr 269
Qy      145 GGGAGTTGGCGA-----PCTCTCCATCTTCCCGCGGACAGAGGGCGGCGC 98
Db      269 oGlyArgArgArgThrGlyAspLeuSerAlaProAlaSer-----ArgSerProAl 286
Qy      97 CGCAGCGCATGGAACTCGAAGTACTTCTCTCGAACTCCCGCGGGTGTCTGGGGGTCTCG 38
Db      286 aAlaLeuArgGlyArgArgAsnGlySerAlaGluSerAlaProGly----- 301
Qy      37 GCCTGCTCTCCGCAATGCGCGCCGCGGT 10
Db      302 -----AsnCysArgArgArg 306

RESULT 86
US-09-252-991A-26078
; Sequence 26078, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26078
; LENGTH: 511
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26078

Alignment Scores:
Pred. No.: 4.14e-07 Length: 511
Score: 165.50 Matches: 83
Percent Similarity: 35.2% Conservatives: 18
Best Local Similarity: 28.9% Mismatches: 117
Query Match: 3.8% Indels: 69
DB: 2 Gaps: 15

US-10-768-158-1 (1-2419) x US-09-252-991A-26078 (1-511)
Qy      720 CCAGCTGTGGCAGTGTCCGTGTCAGGGCTTCAGCTGGGCTTCACAGACACCCCA 661
Db      9 ProAlaGlyAlaLeuAla-----ArgAlaAlaProArgAlaSerValProPro 24
Qy      660 GGAATCTGGCCAGCTGTCTCA-----CCATCGTCACAGGTGCC-----GATGCATGT 613
Db      25 -----ThrProArgAlaProArgSerProProSerProGlyProAlaArgArgProCys 42
Qy      612 CTTCATATCTTGAGAAAAGCAGCTTCGAGT-----CCATCGGTGCTCCAGAACT 562
Db      43 CysAlaValAlaArgProAlaArgAlaGlyArgHisArgProCysArgSerProArgHis 62
Qy      561 CCTGCACGTGCTCAACACGAGGCGGTAGCCAGCTTATCATTCATAAACCCTCGGCGCA 502
Db      63 AlaThrHisValProProGlyAlaArgAlaProAlaArgGlyGlyArgGlyArg 82
```

Qy	501	ATTTCTTGAAGGTGCTCCGTGTT			-----AGCTCA	475
Db	83	ProAspGlyArgCysThrGlyGlyMetalala			ProSerArgThrArgValPheProSer	102
Qy	474	TGTTCCGCAGAGACGGTGGAACTGATAT			TAAGACACCA	415
Db	103	AlaCysArgSerProGlyAspSerGlyAsn			ArgThrGlyAlaArgPro	119
Qy	414	CCATATAGATGACTTGGAGTCTTCATTT			TGGAGGTGCAGCGGCAAAACGGGTAG	355
Db	120	ProAlaProArgProThrSerArgArgAla			ArgAlaArgHisTPrArgLeuAlaAla	139
Qy	354	GGTGGCTTTGTAGCGCGGAGAGGTCA			GTCTCTTGATGATGCCAGGCCCGCGTGTG	295
Db	140	AlaProAlaValArgArgGlySerArg			-----ProAlaGlySerSer	153
Qy	294	GGTACTCCAGGACCGGAGCTCTCGT			CGATGTTCCATCA	238
Db	154	ProArgProGlyProAlaAlaAlaArgGln			ProProValArgArgGlyAlaArgGly	173
Qy	237	CAGCGCTCTGGCTCA			-----CCAGT	217
Db	174	SerAlaProGlySerAlaAspSerArgAla			GlyProPheArgTPrProArgProGly	193
Qy	216	AGACCACCTCTCGACCAAGCTGG			-----TGCCCGACTTGGGGTAGG	175
Db	194	ArgSerProProSerAlaThrAlaArgAla			GlyProValProProLysAlaArgGlyAla	213
Qy	174	TGACGATCCACAGTCGCTGGGCCG			CACCGGAAGTTGGCGATCT	130
Db	214	ProAlaIleProArgGluThrProAlaPro			AlaGlyHisArgAlaAlaArgSerArgAla	233
Qy	129	-----CCTCCATCTTCCCGCGGAGAGG			CGCGCAGCCGCATGGA	79
Db	234	ValArgProProSerAlaAlaArgArgArg			ThrAlaIleProAlaHisAlaPheProArg	253
Qy	78	AGTACTTGTCTCGA			-----ACTCCCCGGGTGCTGGGGGTCTCGG	37
Db	254	SerSerProThrArgAspArgProAlaPro			AlaAlaCysSerGlyCysArgHisLeuAla	273
Qy	36	-----CCTCGCTCTCCGCCA				22
Db	274	HisLeuAlaArgAlaProPro				280

RESULT 87

US-09-252-991A-24321

Sequence 24321, Application US/09252991A

Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252.991A

CURRENT FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 24321

LENGTH: 218

TYPE: PRT

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-24321

Alignment Scores:		
Pred. No.:	3.15e-07	218
Score:	164.50	70
Percent Similarity:	36.6%	Conservative: 19
Best Local Similarity:	28.8%	Mismatches: 83
Query Match:	3.7%	Indels: 71
DB:	2	Gaps: 15

US-10-768-158-1 (1-2419) x US-09-252-991A-24321 (1-218)

771 CAACCTCTTCCCGCCACGGGCA-----GGGCTCAGCGTGTGAGCACTGGTCCACCA 718
Qy |||||
Db GlnAlaSerProAlaProAlaThrGlyAspGlyProAlaArgCysAlaThrGlyAlaPro 32
Qy 717 GCTGGTGGCAGTGCTCGTCAGGCGTTCCA-----GCTGGGCTTGTTCACAGGACACC 664
Db ||||| :|||:|||||
Qy 33 -----GlyAspSerProAlaGlyThrProArgArgSerProAlaArgArgAsnArg 50
Db ||||| :|||:|||||
Qy 663 CCAGGAATCTGG-----CCAGCTGCTCCACCATCGTCACCAAGTCCCGATGCATGT 613
Db ||||| :|||:|||||
Qy 51 ProAlaArgLeuAArgLysAlaProAlaArgProProSerAlaGly-----Ser 67
Db ||||| :|||:|||||
Qy 612 CTTCATACTTGAAGAAAGCAGCTTCGAGTCCATGCGGTGCTCCAGAACTCCTGCACGCT 553
Db :||| :|||:|||||
Qy 68 ArgArgSerArgAlaSerArgAlaGlyLeuCysGly-----AsnArg 82
Db :||| :|||:|||||
Qy 552 GCTCAAAACCCAGGAGCGGTAGCCAGCTTATCATTAACCTCCGGCAGAAATCTTGAA 493
Db ||||| :|||:|||||
Qy 83 AlaArgGlyArgArgGthrCysAlaAspArgAlaThrThrAlaAspSerAlaAsp 102
Db ||||| :|||:|||||
Qy 492 AGGTGCTCGGTAGCTCATGTGCTCGCAGAGCGGTGGAACCTGATAAAGACACCCACCA 433
Db ||||| :|||:|||||
Qy 103 ArgSer---ArgProGlyTyrProAlaAlaGlyProGlyArgProArg----- 117
Db ||||| :|||:|||||
Qy 432 GATCCTTGGGGTGGAGGCCATATAGATGACCTTGGAGTCTCCATTGTGGAGGTGAGG 373
Db ||||| :|||:|||||
Qy 118 -----CysAlaProSerGlyLeuArgTyrProVal----- 127
Db ||||| :|||:|||||
Qy 372 GCAGAAAGCGGTAGCGGAGGTGGCTCTTGATGAGCGGGAGAGGTGAGTCTCTTGATCA 313
Db ||||| :|||:|||||
Qy 128 -----AlaGly-----ArgThrValPro----- 133
Db ||||| :|||:|||||
Qy 312 TGTCCAGGCCGGCTGTGGGTACTCCAGAGACGGGAGCTGCTGCTCGA----- 265
Db ||||| :|||:|||||
Qy 134 LeuAlaGlyArgAlaAlaProValProGlyThrGlyGlySerArgArgCysArg 153
Db ||||| :|||:|||||
Qy 264 -----TGTTCATCAAGCCGATCTCATCGGGTTCAGCGCCT-----GGC 226
Db ||||| :|||:|||||
Qy 154 ArgArgGlyCysAlaAlaAsnProSerCysArgCysArgArgProProAlaAlaCysAla 173
Db ||||| :|||:|||||
Qy 225 TCACCAAGTAGACCACTCTCT----- 205
Db ||||| :|||:|||||
Qy 174 AlaProAlaProProProProHisArgCysArgArgArgSerAlaProAspArgArg 193
Db ||||| :|||:|||||
Qy 204 GCACCAAGCTGGTCCGCGACTTGGGGTAGGTAGCATCCACAGCTCGTGGCGCCGACCG 145
Db ||||| :|||:|||||
Qy 194 HisCysGlyTyrCysArgThrAlaGlyCys--ArgSerAlaSerArgAlaArgThrPro 212
Db ||||| :|||:|||||
Qy 144 GGAAGTTGG 136
Db |||||
Qy 213 ValArgTrp 215
Db |||||

RESULT 88
US-09-252-991A-20553
; Sequence 20553, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEU
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20553
; LENGTH: 315


```
; Sequence 18455, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18455
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18455

Alignment Scores:
Pred. No.: 4,6e-07 Length: 233
Score: 163.00 Matches: 75
Percent Similarity: 36.5% Conservative: 25
Best Local Similarity: 27.4% Mismatches: 92
Query Match: 3.7% Indels: 82
DB: 2 Gaps: 15

US-10-768-158-1 (1-2419) x US-09-252-991A-18455 (1-233)
QY 16 GCGGATCGGAGAGCGAGCGAGCCCGAGCCCGGAGGTTCCGAGGCAAGT 75
Dy : : : : : : : : : : : : : : : : : : : : : : : : : :
2 SerAlaTrpIleGlyLuarG-----AlaThrArgSerSerAlaThrArg 17

QY 76 ACTTCGAGTTCATCGGCTGCGGCTGCGCCCT-----TCTGCGCGGGAAGATGGAGG 129
Dy : : : : : : : : : : : : : : : : : : : : : : : : : :
18 SerSerProIleGluTrpSerGlyLysProCysAlaArgAlaAlaSerLysTrpSer 37

QY 130 AGATCGCCAACTCCCGGTGGCGGCGGAGCGAGTGTGATCGTCACCTACCCCAAGTCCG 189
Dy : : : : : : : : : : : : : : : : : : : : : : : : : :
38 ArgSer-----ValSerProAlaThrIleProArgAlaProCysSerProPro 53

QY 190 -----GCACGAGCTTGCTGAGG-----AGGTGCTTACTTGTGTGAGCCGAGGCG 234
Dy : : : : : : : : : : : : : : : : : : : : : : : : : :
54 AlaAlaArgAlaProAlaCysCysArgSerAlaArgCysSerThrProAlaAlaCysAla 73

QY 235 CTGACCCCGATGAGATCGGCTTGATGAACATCGAGCAGAGCTCCCGGTCTCGAGTACC 294
Dy : : : : : : : : : : : : : : : : : : : : : : : : : :
74 LeuThrTrpSerAlaSerAlaArgAspValAlaAlaAlaGlyCysCysSerAlaSerThr 93

QY 295 CACAGCCGGGCTGGACATCATCAAGGAAGTCACTCTCCCGGCTCATCAAGAGCCACC 354
Dy : : : : : : : : : : : : : : : : : : : : : : : : : :
94 ArgSer-----SerSerSerAlaPro 100

QY 355 TGCCCTACCGCTTTCTGCCCTCTG-----ACCTCCACA 387
Dy : : : : : : : : : : : : : : : : : : : : : : : : : :
101 CysProSerThrTyArgProThrIleValProSerProTrpProMetGlyThrSerGly 120

QY 388 ATGAGAGACTCAAGG-----TCATCTATATGGTCTC-----GCAACCCCAAGGATCGG 435
Dy : : : : : : : : : : : : : : : : : : : : : : : : : :
121 CysSerAlaProLysGlyTrpAlaSerSerIleAlaIleProSerSerAlaSerAlaTrp 140

QY 436 TGGTGTCTTATATATCAGTTCACCGCTCTTCGCGGACCATGATGCTACCGAGCCCTTTC 495
Dy : : : : : : : : : : : : : : : : : : : : : : : : : :
141 ArgCysThrAsnMetAlaGlyThrCysTrpAsnMetProSerIleThrSerVal----- 158

QY 496 AAGAATTCTGCGGAGGTTTATGAATATAAGCTGCGGTACGGCTCTCTGTTTTCGAGCAG 555
Dy : : : : : : : : : : : : : : : : : : : : : : : : : :
159 -----AlaThrGlyGlyArg-----ArgAlaAlaProGlyAlaSerAsn 171

QY 556 TGCAGGAGTCTTGGGAGCACCAGTGGAGTCAAGCTGCTTTCTCAAGTATGAAGACA 615
Dy : : : : : : : : : : : : : : : : : : : : : : : : : :
172 -----AlaAlaAlaArgThr 176
```

```
QY 616 TGCATCGGACCTGGTGACGATGGTGAGCAGCTGCCAGATTCCTCGGGGTGCTGTG 675
Dy : : : : : : : : : : : : : : : : : : : : : : : : : :
177 CysSerAlaArgTrp-----ProTrpArgProAlaTrpGlyCys----- 189

QY 676 ACAAGGCCAGCTGGAAGCCCTGACGAGCACTGCCACCACTGGTGACAGTGTGCA 735
Dy : : : : : : : : : : : : : : : : : : : : : : : : : :
190 -----TrpArg-----ArgSerAlaTrpSerGlySerAla 199

QY 736 ACGTGTAGGCGCTGCCCGTGGCGCGGGAAGAGTTGGGTGT 777
Dy : : : : : : : : : : : : : : : : : : : : : : : : : :
200 GluThrSerProSerAlaSerSerGlyTyArgThrValCys 213

RESULT 92
US-09-252-991A-16586
; Sequence 16586, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 16586
; LENGTH: 246
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-16586
```

```
Alignment Scores:
Pred. No.: 4,75e-07 Length: 246
Score: 163.00 Matches: 77
Percent Similarity: 34.8% Conservative: 19
Best Local Similarity: 27.9% Mismatches: 92
Query Match: 3.7% Indels: 88
DB: 2 Gaps: 13

US-10-768-158-1 (1-2419) x US-09-252-991A-16586 (1-246)
QY 762 CCGGCCCCAGCGAGCGGCTCAGCTTGCGAGCACTGTCACCACTGGTGGCAGTGCT 703
Dy : : : : : : : : : : : : : : : : : : : : : : : : : :
2 ProProGluArgTyGlnValGluArgAlaGlyThrGlyTyProAlaGlyGlySerAsp 21

QY 702 CCGTCA-----GGCTTCCAGCTGGG 682
Dy : : : : : : : : : : : : : : : : : : : : : : : : : :
22 ProAlaHisArgArgArgSerProArgGluArgArgArgGlyIleProAlaAla 41

QY 681 CTTGTGCACAGGACACCCCGAGGAATCTGCCAGCTCTCCACCATCGTCACCAAGTCCC 622
Dy : : : : : : : : : : : : : : : : : : : : : : : : : :
42 ProSerHisArgArgArgArg-----ProLeuArgHisProArgSerGluGly--- 57

QY 621 GATGCATGCTCTCATCTTCTGAGAAAAGCAGCTTCGAGTCCATCGCGTGTCCAGAACT 562
Dy : : : : : : : : : : : : : : : : : : : : : : : : : :
58 -----AspArgArgArgAla-----GlyArgSerArgArg 67

QY 561 CTGTCACGTGCTCAACACGAGCGCTAGCGAGCTTATCATTAACCTCCGCGAGA 502
Dy : : : : : : : : : : : : : : : : : : : : : : : : : :
68 ArgSerGlySerSerThrProGlnArgCysPro----- 78

QY 501 ATTCTTGAAGAGTGCCTCGGTAGTCTATGTTCCGACAGAGCGGTGGAACCTGATAAAG 442
Dy : : : : : : : : : : : : : : : : : : : : : : : : : :
79 -----CysAlaArgArgSerArgSerAla-----GlyArgGlyArgArg 91

QY 441 ACACCACAGATCTTGGGGTTGGAGCCATATAGATGACCTTGGAGTCTCCATTGTGGA 382
Dy : : : : : : : : : : : : : : : : : : : : : : : : : :
92 ArgProProAsnPro-----SerProArgArgHisArgGlnProArgArgSerArg 108
```



```
US-10-768-158-1 (1-2419) x US-09-252-991A-30843 (1-663)
Qy 2133 AGTCCTCAGTAAACATCCATACAGTACTAAATAATAGAAAAATATAAATGAAATTTTCACAGA 2254
Dy 18 AlaAlaArgAlaSerProTrpSerThrThrArg
Qy 2253 AAGCAGCTCCCTCCACAGAAACACAGGCAAGGCTGCTGCCGAGCCACCATCGCATGCCAC 2194
Dy 30 -----ProProAlaAseP-----ArgSerGlyAlaThrAlaArgValProAlaCysArg 45
Qy 2193 GTGCCACT-----GCCTGG----- 2179
Dy 46 ArgValProIleAArgCysAlaAsnAlaTrpProGlyArgProMetSerSerThrSerArg 65
Qy 2178 ---CCAGCTTCTCCCATTAACAACAACAAAAACACAGCTTCTCTACACCGCTTTCAAA 2122
Dy 66 ArgProGlyTrpSerProAlaValProAlaGlyAlaCysCysTrpLysThrArgAsnVal 85
Qy 2121 TACAGCAACAAAGACGGAGACCCCTTCCACAGCAGCGACCCCTCTGACCACTTTGGG 2062
Dy 86 AlaGlyArgSerAlaThrAspAlaTrpSerSerAlaProAlaProAlaAsnCysCys 105
Qy 2061 TGCTTTCAGCCACATAAATGACAGAACACCCGCTCCCAAGGGGTTGGCTGCTACTG 2002
Dy 106 LeuPheProAla-----GlyProSerPro----- 113
Qy 2001 GCCCCGACTCTCATCGCTTCAGGAAATGGGTGAGAAA----- 1960
Dy 114 AlaSerProAlaProAlaAlaCysAArgArgTrpProArgAlaAlaCysHisTrpProAla 133
Qy 1959 -----TTAGGCGATGAAGCAATAAAGACCTCACCTTTACTCCCTCAG 1918
Dy 134 SerAlaTrpTrpTrpLeuAlaProAlaAlaArgCysCysTrpProAlaProAlaProAla 153
Qy 1917 ATACACGGA-----GGATGAGGTGGACAGACGCGCGCGCTGCCAAC 1870
Dy 154 SerAlaGlyAArgAlaCysCysAlaSer-ProAsnAArgArgArgGluProTrpProPr 173
Qy 1869 TGATCCG-----AGTGAAACAGGCTACACTCGCAAAATGGTCTCC----- 1827
Dy 173 oSerProTrpAlaSerArgAlaGlyProAlaSerCysGlyArgProProAlaCysSerPr 193
Qy 1826 -----ACGGCTGAGCGCTCAGGTGCACAGAACAACTGAATGATAAAGACCGGGA 1774
Dy 193 oValAlaThrAlaProThrAlaThrCysSer-----ProProSe 206
Qy 1773 TTTTGAATAATTACTATTAAGCATCTGCTGCCAGCATTAAGGCTATCATCTTCAGAA 1714
Dy 206 rAlaArgSerAlaTrpLysProCysAlaCysAlaLysAlaValGly----- 221
Qy 1713 ACAAGTAAAGCTGTGGAATTCAGCTCAATTCGAGGCGCACTCGGTCTTTGGTGCC 1654
Dy 222 -----SerAlaArgSerProAlaAsnAlaTrpProAlaAlaSerAlaTrpCysPr 238
Qy 1653 GACTCTCACCCACTGCATGCAATGCTTCCTGACTGCTAGCTCCACCTGCCAGCT 1594
Dy 238 oThrCysAsnSerAlaArgProSerAla-lleAlaSerThrAlaArgArgSerProSer 258
Qy 1593 CACGAGGAGCA-----CTTCCCTCAGCATGCTCTCCGCGGCTC 1555
Dy 258 hrAenGlyArgProAlaAlaCysProThrThrTrpProProAlaAsnAlaPro----- 275
Qy 1554 CTCTGCATTTGGGCTTTGGAGCCGCTGGAAGTCTCTGGCAGAGGAGGCTCATCT 1495
Dy 276 -----AlaSerAlaAlaAlaAsnTrp-ProTrpTrpLysAlaProSerProAla 291
Qy 1494 GAAGACCTAGTGAAACAATTTCTGCTCATTTCCAGGAAGGTTCTGAAGAAGTGAACCTCGCTC 1435
Dy 291 ----- 291
Qy 1434 AGAATAAACAGTGTCCAGATTCCCCCAACAGACCCCAACTCATCTCTGTTCTCACAC 1375
Dy 292 -----MetProProSerThrSerAlaThrArgLeuAlaAlaCysGlyHis--- 306
```

```
1374 TATCTCAGAACCAAGTGAGAACAGAAATGGCTGTGCCTTATATAAACTACAAAATA 1315
Dy 307 -----AspGlyAseAlaGlyArg--- 312
Qy 1314 AAGAGTTAAACCTCAGATCTTTCTACCAATTCGGGTGGCTCGCTCTCGATTCCTCCCTGG 1255
Dy 313 -----AlaSerGlnThrProTrp 318
Qy 1254 AAATGAACCTTTATTTGGTTTACTGACATTTATGTAGATTTCAGTGAAAAAGCTCTATAA 1195
Dy 319 ArgAspThr----- 321
Qy 1194 AATACATAATACGGGTTCGAAAGGACAGACATTCTAGTTGCATA----- 1150
Dy 322 -----SerProCysValProAsnCysValAsn 330
Qy 1149 TATTACAGGCTTTATCCTTACGCTCCAGGCAATTTGGAACATGTGGAGACTGTTTGT 1090
Dy 331 TrpArgArgArgIlePro-----TrpSerAlaValAlaLysThrSerPro 345
Qy 1089 AATCAGACATGGAGAGGCTCCAGCTTCTAAAGGGGAGACAGCTCTTCGGTTGGGAATCA 1030
Dy 346 TrpProArgTrpProGlyMetProAlaGlyProArgProSerCysThrAlaValAlaAla 365
Qy 1029 TCACACTCCCTCCGCTCACGCCCTCTTCCCTTCCCGCTGTTTCACAGCTGCTTCCA 970
Dy 366 TrpAlaProAlaArgGlyGlyTyAlaAlaAlaProArg----- 378
Qy 969 GAGTTTCCAGCAAGGAATAATGAATGCATACAGACATTTTGGCTAGTAGCTGCTG 910
Dy 379 -----AsnSerCys-----SerAla 383
Qy 909 GGTATTGTGACATGCAGGTTGTTGTTCTGTTATTATAAATAAAGTCAAACTGAGGT 850
Dy 384 GlyArgLeuArgArgArgValArgArgSer-----AlaArgAlaTrp 398
Qy 849 CACACTTTCCTCCTCTTATACACCAAGTCAAACTTCTCAT-----TCATGGAGA 796
Dy 399 ArgProTrpProAlaGlyArgAlaThrProAlaSerAspAlaArgArgAseProAlaSer 418
Qy 795 CGGTGAAGATGTCTTCCACAGCCCAACTCTTCCCGGCCACCGGCGAGGCGCTCAGCGT 736
Dy 419 ArgGlyArg---ProAlaThrAlaAlaAspHisProAlaAlaTrpValSerAlaAlaArg 437
Qy 735 TGACAGCACTGGTCCACAGCTGGTGGCAGTGTCCGTCAGGGCTTCACAGCTGGCCCTGT 676
Dy 438 ArgThrSerSerAlaProIleAlaGlySerAlaProGlySerGlyThrAlaProArgCys 457
Qy 675 CACAGGACACCCCGAGAAATCTGGCCAGCTCTCCACCATCGTCACCCAGGTCGCCGATGCA 616
Dy 458 His-----ProValArgLysAseGlyAlaGlyProAlaSerThrGly----- 471
Qy 615 TGTCTTCATCTTGAGAAAAAGCAGCTTCGAGTCCATCGGTCTCCAGAGCTCTCCAGAACTCTGCA 556
Dy 472 -----ArgSerArgArgArgTrpAlaAlaAlaProAlaAlaArgAla 485
Qy 555 CGTGCTCAACACGAGGAGCCGTAGCCAGCTTATCATTTCAAACTCCGCGAGAACTCTT 496
Dy 486 ArgThrGlyIleArgArgThrAlaAlaAla----- 495
Qy 495 GAAAGGTGCCTCGTAGCTCATGTCGCCAGAGAGCGGTGGAACGTAGTAATAAGACACCA 436
Dy 496 ---AlaCysTyr-----ProPro 500
Qy 435 CCAGATCTTTGGGTTTGGCGCCATATAGATGACCTTTGGAGTCTCCATTGTGGAGGTGAC 376
Dy 501 ProAlaPro-----GlyArg 505
Qy 375 AGGCAGAAAGCGGTAGGCGAGGTGCTTGTATGAGCGGGGAGAGCTCAGTCTCTTGA 316
Dy 506 ArgThrProAlaAlaArgArgArgSerAlaArgThrAlaProArgProAlaIleProAla 525
```

```
Qy 315 TGATGTCCA---GGCCCGCTGTGGTACTCCAGAGCCGGAGCTGCTCGTGGATTTCA 259
Db      |||
526 HisArgProSerAlaSerAlaThrGlyTrpProGlyProAlaLeuAla-----CysPro 543
Qy      |||
258 TCAAGCGATCTCATCGGGTCCAGCGCTCGCTCACCAGTAGACACCACTCTCTGCAGCA 199
Db      |||
544 AlaAla-----AlaGlyArgProSerProAlaProAlaProAlaPro 560
Qy      |||
198 AGCTGGTCCGGAAGTGGGGTAGGTGACGATCCACACGCTCGCTGGCGGCACCGGAAGT 139
Db      |||
561 His-----AlaArgProArgAlaProArgArg-----ArgProGly--- 572
Qy      |||
138 TGGGATCTCTCATCTTCCCGGGCAGAGGGCGGAGCGGCGCAGCCATGGAACTCGA 79
Db      |||
573 ---ArgSerProAlaThrAlaProAlaAlaPheAlaAlaProAla----- 586
Qy      |||
78 AGTACTTGCTCTCGAACTCCCGGGGTCTGGGGTCTCGGCTCGCTCTCCGCCATGC 19
Db      |||
587 GlyProThrProAlaArgProProAlaAlaAlaThrGlyArgProSerAlaProProArg 606
Qy      |||
18 CGCCGCTCGCCGT 4
Db      |||
607 ProAlaArgArgArg 611

RESULT 95
US-09-252-991A-26899
; Sequence 26899, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26899
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26899

Alignment Scores:
Pred. No.: 6.7e-07 Length: 247
Score: 161.50 Matches: 72
Percent Similarity: 28.3% Conservative: 12
Best Local Similarity: 24.2% Mismatches: 88
Query Match: 3.7% Indels: 125
DB: 2 Gaps: 15

US-10-768-158-1 (1-2419) x US-09-252-991A-26899 (1-247)
Qy 786 TGCTCTCCACAGCCCACTCTCCCGGCCAGCGGCGGCTCGAGCTTCAGCACT 727
Db      |||
2 CysProSerCysSerPro-----ProSerPro---CysSerProLeuArgArgTrpPro 18
Qy 726 GGTCCACAGCTGTGGCAGTGTCCGTCAGGGCTTCAGCTGGGCT----- 679
Db      |||
19 SerProThrProAlaThrAlaAsnThrAlaTrpSerProAlaSerProIleArgSerPro 38
Qy 678 -----TGTACAGGACACCCCGGAGTCTGGCCAGTCTGGCTGCTCCACCATCTCA 631
Db      |||
39 AlaSerThrThrCysTrpArgCysSerProSer-----AlaSerGlyProProSerAsn 56
Qy 630 CCAGTCCCGATGCATGCTTTCATCTTACACTTGAGAAAAAGCAGCTTCGAGTCCATCGGTGCT 571
Db      |||
57 ArgGluPro-----ArgValSerProCysProAla 66
Qy 570 CCCAGAACTCTGTCAGCTGCTCAAAACCAGGAGCGGTAGCCCGCTTATCATTTATAAAC 511
```

```
Db 67 Pro-----SerAlaProCysTrpSerAlaGlyCys 77
Qy 510 TCCGGCAGAAATCTTGAAGGTGCTCGTAGCTCATGCTCCGAGAGCGGTGGAACT 451
Db      |||
78 SerAlaSer-----ArgAlaCysAsnCysArgSerTrp----- 88
Qy 450 GATAAATAAGACACACAGATCTTGGGGTGGAGCCATATAGATGAACCTTGAGTCTC 391
Db      |||
89 -----IleProAlaSerProProArgCysSerProSerAlaSerAlaTrpProSer 105
Qy 390 CATTGT----- 385
Db      |||
106 ProCysAlaArgArgCysArgTrpArgTrpProProArgCysSerProTrpProThr 125
Qy 384 GGAGGTCCAGAGGGCAGAAAGCGGTAGGCGCAGGTGGCTCTTGATGAGGCGGGAGGTCA 325
Db      |||
126 GlySerProThrAlaTrpSerCysArgThr----- 135
Qy 324 GTTCTTGTGATGATCCAGCCCGGCTGTGGGTACTCCAGAGCCGGAGTGTCTGTCGA 265
Db      |||
136 -----SerProAlaProGlyCysThrArgSer 144
Qy 264 TGTTCATCAAGCCGATCTCATCGGGGTACAGGCGCTGGCTCACCAGTAGACCACTCCT 205
Db      |||
145 ValSerSer-----ProProProProArgCysThrProPro 156
Qy 204 GCAGAAAGTGTGCGGACTTGGGGTAGGTAGTACACATCCACACGCTCGCTG---GCCGCA 148
Db      |||
157 AlaThrArgTrpCysAlaSerTrpArg-----ArgArgProArgProTrpTrpValSer 174
Qy 147 CCGGGAAGTTGGGATCTCTCATCTTCCCGGGCAGAGGGCGGCGGACGCCAT 88
Db      |||
175 ProGlyTrpProArgArgAlaProGlySer----- 184
Qy 87 GGAAGTCAAGTACTTCTCGAACTCCCGGGGTGCTGGGGTCTCGGCT----- 34
Db      |||
185 -----GlyCysTrpProAlaGluProProArg 193
Qy 33 -----CGCTCTCCGCAATCGCGGCTCGCTCGCT 7
Db      |||
194 LeuProGlnArgLeuIleSerThrArgArgProTrpCysThrArgArg 210

RESULT 96
US-09-252-991A-31371
; Sequence 31371, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31371
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31371

Alignment Scores:
Pred. No.: 7.09e-07 Length: 272
Score: 161.50 Matches: 84
Percent Similarity: 36.5% Conservative: 13
Best Local Similarity: 31.6% Mismatches: 87
Query Match: 3.7% Indels: 82
DB: 2 Gaps: 11
```

US-10-768-158-1 (1-2419) x US-09-252-991A-31371 (1-272)

```
Qy 648 GCTGCTCACCATCGTCAACGAGTCCCGATGCATGCTTTCATCTTGTAGAAAAACGACGT 589
      |||||
Db 24 AlaAlaThrGlyValArgProGlyPro-----HisArgGlyAlaArgGln 39
      |||||
Qy 588 TCGAGTCCATGC---GGTGTCCAGAACTCCTGCAGTGTCTCAACACGAGCGGTAGC 532
      |||||
Db 40 SerAlaGlyCysProGlyAlaProArgProGlyArg----- 52
      |||||
Qy 531 CCAGCTTATCATTAACCTCCGGCAGAAATCTTGAAGGTGCTCGGTAGTCTATGG 472
      |||||
Db 53 ---Ala-AlaIleTrpArgProGlyAlaArgGlyArgProProAlaHisArg 71
      |||||
Qy 471 TCCGAGAGAGCGGTGGAACTGATAATAAGACACCACCATGATCCTTGGGTGGAGCCA 412
      |||||
Db 71 gArgValArgAlaArgPheArgSerArgSerArgHisArgAlaGlyValAla-Ala- 90
      |||||
Qy 411 TATAGATGACCTTGGAGTCTCCATGTGTGGAGTCCAGGGCGAGAAAGCGGTAGGCAAGT 352
      |||||
Db 91 -----GlyArgCysArg-----ArgL 96
      |||||
Qy 351 GGCTCTTGATGAGCGGGAGAGGTCACTTCTTGATGATGTCCAGGCCCGCTGTGGT 292
      |||||
Db 96 euAlaGlnLeuArgProGlyProValaPheAlaAlaLeuAlaArgProGlyCys-Arg 115
      |||||
Qy 291 ACTCAGAGCCGGAGCTGCTGCTGATGTTTCATCAAGCCGATCTCATCGGGGTACGGC 232
      |||||
Db 116 ArgArgGlyProAlaAlaSerArgGlyCysAlaGlyAlaGly-----Leu 131
      |||||
Qy 231 CCTGCTCACCAGT-----AGACCACT-----CCTGCAGCA 199
      |||||
Db 132 ProGlyGlyProSerAlaProCysArgAlaProGlnArgProGlyArgProAlaLeu 151
      |||||
Qy 198 AGCTGGTCCCGACTTGGGGTAGTGACGATCCACACGCTCGCTGG----- 154
      |||||
Db 152 ProValAlaArgProArgArgArgSerArgArgArgThrGlyProValaArg 171
      |||||
Qy 154 ----- 154
      |||||
Db 172 ThrAlaGlyThrAlaArgArgArgProGlyAlaSerProAsnProGlyAla 191
      |||||
Qy 153 -----GCCGACCGGGAAGTTGGCGATCTCTCCATCTTCCCGC 115
      |||||
Db 192 GlyValArgProArgProAlaAlaGlyArgGlySerArgArgLeuProArgHisAlaArg 211
      |||||
Qy 114 GGC-----AGAAGGGCGCAGCGCCATGCCACTCGAAGTACTTGCTCTCGAACT 61
      |||||
Db 212 GlyArgValArgArgArgAlaAlaProGlyThrArgSerHisArgThrAlaGly 231
      |||||
Qy 60 CCCCCGGGTGCTGGGGTCTCGG-----CCTCGCTCTCCGCCATGCCGCGC 13
      |||||
Db 232 ArgProAlaCysArgSerArgGlyLeuLeuLeuProArgLeuProProAlaArg 251
      |||||
Qy 12 CGTCCCGCTGC 1
      |||||
Db 252 AlaArgValArg 255
```

RESULT 97

US-09-252-991A-28132
; Sequence 28132, Application US/09252991A
; Patent No. 6551795

GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.

FILE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 28132

LENGTH: 333

TYPE: PRT

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-28132

Alignment Scores:

Pred. No.:	1e-06	Length:	333
Score:	160.50	Matches:	64
Percent Similarity:	35.5%	Conservative:	7
Best Local Similarity:	32.0%	Mismatches:	54
Query Match:	3.6%	Indels:	75
DB:	2	Gaps:	11

US-10-768-158-1 (1-2419) x US-09-252-991A-28132 (1-333)

```
Qy 438 CCACGATCCTTGG-----GGTTGCCAGCCATATAGATGACCTTGG 397
      |||||
Db 2 ProProAlaProArgProAlaAlaArgArgProGlyCysAlaGlyArg- 18
      |||||
Qy 396 AGTCTCCATTGTGAGGTAGAGGCGAGAAAGCGGTAGGCGAGTGTGCTCTTGATGAGC 337
      |||||
Db 19 -----GlyArgSerArgProGlyArgArgGlyArgCysSerAla 33
      |||||
Qy 336 GGGAGAGGTTCAGTTCCTTGTATGATGTCAGGC-----CCGGCTGGGTACTCCA 286
      |||||
Db 34 GlySerArg-----ProGlyTrpArgTrpProAlaAlaThrAla 47
      |||||
Qy 285 GGACCGGAGTCTGCTCGATGATGTTATCA----- 256
      |||||
Db 48 GlyArgCysAlaGlyArgArgCysGlyArgSerArgThrAlaProSerSerIleGly 67
      |||||
Qy 255 -----AGCCGATCTCATCGGGGTGAG 235
      |||||
Db 68 GlyTrpArgHisAlaGlyGlyAspGlyArgArgArgGlyArgGlyHisArgAspAsn 87
      |||||
Qy 234 -----CGCCCTGGCTCACC----- 220
      |||||
Db 88 ProCysArgCysSerArgAspTrpArgArgAspGlyArgProGlyAlaAlaGlyArg 107
      |||||
Qy 219 -----AGTAGACACCTCTTCGACCAAGCTGTGTCGG-----ACTTGGGGT 178
      |||||
Db 108 ThrCysArgSerArgProGlyProProTrpHisTrpProArgArgGlySerValGly 127
      |||||
Qy 177 AGGTGACATCCACATCCTGCTGGCGCCACCGGGAAGTTGGCGATCTCTCCATCTTCC 118
      |||||
Db 128 ArgGlyArgGlySerAlaArgTrpLeuProProGlySerAlaArgSerValAlaAlaArg 147
      |||||
Qy 117 CGCGGCGAGGGCGGCGCGCCGACCGCATCGAACTCGAAGTACTTCTCTCGAACTCCC 58
      |||||
Db 148 ArgGly---ArgAlaAsnArgCysArgGly-----ArgArgPro 160
      |||||
Qy 57 CCGGGGTGCTGG---GGGTCTCGGCTCTGCTCTCGCATCGCCCGCGCTCGCGCTGC 1
      |||||
Db 161 ProGlyArgArgProGlyArgArgCysArgThrArgProAlaAlaArgArgArgArg 180
      |||||
```

RESULT 98

US-09-252-991A-30843

Sequence 30843, Application US/09252991A
; Patent No. 6551795

GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.

FILE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142

```
; SEQ ID NO 30843
; LENGTH: 663
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30843

Alignment Scores:
Pred. No.:      1.51e-06      Length:      663
Score:          160.50      Matches:      93
Percent Similarity: 33.7%      Conservative: 27
Best Local Similarity: 26.1%      Mismatches: 111
Query Match:      3.6%      Indels:      125
DB:              2          Gaps:      22

US-10-768-158-1 (1-2419) x US-09-252-991A-30843 (1-663)
QY 1 GCGACGGCGACGGCGCGCATGGCGGAGA-----GCGAGGCGG 39
Db 1 GCGACGGCGACGGCGCGCATGGCGGAGA-----GCGAGGCGG 39
QY 37 AlaThrAlaArgValProAlaCysArgValProIleArgCysAlaAsnAlaTrpPro 56
Db 37 AlaThrAlaArgValProAlaCysArgValProIleArgCysAlaAsnAlaTrpPro 56
QY 40 AGACCCCGACGACCCCGGGGAGTTCGAGACGCAAGTACTTCGAGTTCCATGCGCGGCG 99
Db 40 AGACCCCGACGACCCCGGGGAGTTCGAGACGCAAGTACTTCGAGTTCCATGCGCGGCG 99
QY 57 GlyArgProMetSerThrSerArgArgProGlyTrpSerProAlaValProAlaGly 76
Db 57 GlyArgProMetSerThrSerArgArgProGlyTrpSerProAlaValProAlaGly 76
QY 100 TGC-----CGCCCTCTCCCGCGGGAAGA-----TGGAGGAGA 132
Db 100 TGC-----CGCCCTCTCCCGCGGGAAGA-----TGGAGGAGA 132
QY 77 AlaCysCysTrpLysThrArgAsnValAlaGlyArgSerAlaThrAspAlaTrpSerSer 96
Db 77 AlaCysCysTrpLysThrArgAsnValAlaGlyArgSerAlaThrAspAlaTrpSerSer 96
QY 133 TCGCCAACCTCCCGGTGCGGCCGACGACGTGT-----GGATCGTCA 174
Db 133 TCGCCAACCTCCCGGTGCGGCCGACGACGTGT-----GGATCGTCA 174
QY 97 AlaPro-----AlaProAlaAsnCysCysLeuPheProAlaGlyProSer 112
Db 97 AlaPro-----AlaProAlaAsnCysCysLeuPheProAlaGlyProSer 112
QY 175 CCTACCCCAAGTCCGCGACCACTTGCTGCGAGGAGTGGTCTACT-----TGG 222
Db 175 CCTACCCCAAGTCCGCGACCACTTGCTGCGAGGAGTGGTCTACT-----TGG 222
QY 113 ProAla---SerProAlaProAlaAlaCysArgArgTrpProArgAlaAlaCysHisTrp 131
Db 113 ProAla---SerProAlaProAlaAlaCysArgArgTrpProArgAlaAlaCysHisTrp 131
QY 223 TGAGCCAGGGCG-----CTGACCCCGATGAGATCGGCTTGATGAACATCGACGAGC 273
Db 223 TGAGCCAGGGCG-----CTGACCCCGATGAGATCGGCTTGATGAACATCGACGAGC 273
QY 132 ProAlaSerAlaAlaTrpTrpTrpLeuAlaProAlaArgCysCys-----145
Db 132 ProAlaSerAlaAlaTrpTrpTrpLeuAlaProAlaArgCysCys-----145
QY 274 AGCTCCCGGTCTCTG-----AGTACCCACAGCGCGGCTGGACATCA 315
Db 274 AGCTCCCGGTCTCTG-----AGTACCCACAGCGCGGCTGGACATCA 315
QY 146 -----TrpProAlaProProAlaProAlaSerAlaGlyArgAlaCysCysAla 161
Db 146 -----TrpProAlaProProAlaProAlaSerAlaGlyArgAlaCysCysAla 161
QY 316 TCAAGGAAC-----TCACCTCTCCCGCGCTCATCAAGA-----348
Db 316 TCAAGGAAC-----TCACCTCTCCCGCGCTCATCAAGA-----348
QY 162 SerProAsnArgArgArgGluProTrpProProSerProTyTrpAlaSerArgAlaGly 181
Db 162 SerProAsnArgArgArgGluProTrpProProSerProTyTrpAlaSerArgAlaGly 181
QY 349 ---GCCACTGC-----CCTACCGCTTCTGCGCTCTGACCTCACAATGGAGACTCCA 399
Db 349 ---GCCACTGC-----CCTACCGCTTCTGCGCTCTGACCTCACAATGGAGACTCCA 399
QY 182 ProAlaSerCysGlyArgProProAlaCysSerProValAlaThr-----196
Db 182 ProAlaSerCysGlyArgProProAlaCysSerProValAlaThr-----196
QY 400 AGGTCACTATATGCTCGCAACCCCAAGGATCTGGTGGTGTCTTATTATTATTCACATTCCACC 459
Db 400 AGGTCACTATATGCTCGCAACCCCAAGGATCTGGTGGTGTCTTATTATTATTCACATTCCACC 459
QY 197 -----AlaProThr 199
Db 197 -----AlaProThr 199
QY 460 GCTCTCTCGGACCATGAGTACCGAGGCACCTTTCAAGAATTCTGCGGAGGTTTATGA 519
Db 460 GCTCTCTCGGACCATGAGTACCGAGGCACCTTTCAAGAATTCTGCGGAGGTTTATGA 519
QY 200 AlaThrCysSerProProSerAlaArgSerAlaTrpLys-----212
Db 200 AlaThrCysSerProProSerAlaArgSerAlaTrpLys-----212
QY 520 ATGATAAGCTGGGTACGGCTCTGGTTGGTTGAGCAGCTGCAGAGT-----564
Db 520 ATGATAAGCTGGGTACGGCTCTGGTTGGTTGAGCAGCTGCAGAGT-----564
QY 213 -----ProCysAlaCysAlaLysAlaValGlySerAlaAlaArgSerProAlaAsnAlaTrp 230
Db 213 -----ProCysAlaCysAlaLysAlaValGlySerAlaAlaArgSerProAlaAsnAlaTrp 230
QY 565 TCTGGGACCGCATGAGTACGACGTGCTTTTCTCAAGTATGAAGACATGATCGGG 624
Db 565 TCTGGGACCGCATGAGTACGACGTGCTTTTCTCAAGTATGAAGACATGATCGGG 624
QY 231 ProAlaAlaSerAlaTrpCysProThrCysAsnSerAlaArgProSerAlaAlaSer 250
Db 231 ProAlaAlaSerAlaTrpCysProThrCysAsnSerAlaArgProSerAlaAlaSer 250
QY 625 ACCTGGTGACCATGCTGAGTACGACGTGCTTTTCTCAAGTATGAAGACATGATCGGG 684
Db 625 ACCTGGTGACCATGCTGAGTACGACGTGCTTTTCTCAAGTATGAAGACATGATCGGG 684
QY 251 ThrAlaAlaArgSerProSerThrAsnGlyArgProAlaAlaCysProThrThrPro 270
Db 251 ThrAlaAlaArgSerProSerThrAsnGlyArgProAlaAlaCysProThrThrPro 270
QY 685 AGCTGGAAGCCCTGACGGGACGACTGCCACCAGC-----TGGTGG-----723
Db 685 AGCTGGAAGCCCTGACGGGACGACTGCCACCAGC-----TGGTGG-----723

RESULT 99
US-09-252-991A-32166
; Sequence 32166, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32166
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32166

Alignment Scores:
Pred. No.:      8.51e-07      Length:      208
Score:          160.00      Matches:      77
Percent Similarity: 38.5%      Conservative: 22
Best Local Similarity: 30.0%      Mismatches: 90
Query Match:      3.6%      Indels:      68
DB:              2          Gaps:      14

US-10-768-158-1 (1-2419) x US-09-252-991A-32166 (1-208)
QY 10 ACGCGGGCGGCGGCGGAGCGAGCGGCGGAGACCCCGGCGCA-----CCCGGGGGAGT 63
Db 10 ACGCGGGCGGCGGCGGAGCGGAGCGGCGGAGACCCCGGCGCA-----CCCGGGGGAGT 63
QY 1 SerSerIleAlaArgArgArgThrProProLysCysArgSerTrpTrpProArgSerSer 20
Db 1 SerSerIleAlaArgArgArgThrProProLysCysArgSerTrpTrpProArgSerSer 20
QY 64 TCGAGAGCAAGTACTTCGAGTTCCATGCGGTGCGGCTGCGGCCCTTCTGCGCGGGAAGA 123
Db 64 TCGAGAGCAAGTACTTCGAGTTCCATGCGGTGCGGCTGCGGCCCTTCTGCGCGGGAAGA 123
QY 21 GlyPheProSerThrAlaSerSerSerThrCysAlaAlaTrpAlaAlaSerAla---39
Db 21 GlyPheProSerThrAlaSerSerSerThrCysAlaAlaTrpAlaAlaSerAla---39
QY 124 TGGAGGAGATCGCCCAACTTCCCGGTGCGGCGGCGGAGCGAGCTGTGGATCGTCACT---ACC 180
Db 124 TGGAGGAGATCGCCCAACTTCCCGGTGCGGCGGCGGAGCGAGCTGTGGATCGTCACT---ACC 180
QY 40 --AlaArgLysProArgProArgArgArgProAlaCysAlaArgTrpSerProThrThr 58
Db 40 --AlaArgLysProArgProArgArgArgProAlaCysAlaArgTrpSerProThrThr 58
QY 181 CCAAGTCCGGCACCAGCTTGC---TGCAGGAGGTGGTCTACTTGGTGAGCGGCGGCTG 237
Db 181 CCAAGTCCGGCACCAGCTTGC---TGCAGGAGGTGGTCTACTTGGTGAGCGGCGGCTG 237
QY 59 ProGlyValArgArgArgCysAlaCysArgAlaTrpArgThrCysArgSerProAlaSer 78
Db 59 ProGlyValArgArgArgCysAlaCysArgAlaTrpArgThrCysArgSerProAlaSer 78
QY 238 ACCCGATGAGATCGGCTTGATGAACATCGACGAGCAGCTCCCGGTCTGGAGTACCCAC 297
Db 238 ACCCGATGAGATCGGCTTGATGAACATCGACGAGCAGCTCCCGGTCTGGAGTACCCAC 297
QY 79 ValThr-----ArgSerThrSerAsnThrThrSerAlaSerThrThr 92
Db 79 ValThr-----ArgSerThrSerAsnThrThrSerAlaSerThrThr 92
QY 298 AGCGGGCGCTGGACATCATCAAGGAAC---TGACCTCTCCCGGCTCATCAAGAGCCACC 354
Db 298 AGCGGGCGCTGGACATCATCAAGGAAC---TGACCTCTCCCGGCTCATCAAGAGCCACC 354
QY 93 MetAlaAlaCysThrAlaSerArgSerThrTrpProAlaIleAlaIleArgArgThr 112
Db 93 MetAlaAlaCysThrAlaSerArgSerThrTrpProAlaIleAlaIleArgArgThr 112
QY 355 TGCCCT-----ACGGCTTCTGCGCTCTGACCTCCACAAATGGAGACTCCAGG 402
Db 355 TGCCCT-----ACGGCTTCTGCGCTCTGACCTCCACAAATGGAGACTCCAGG 402
```

```
Db 113 SerProAlaArgSerSerThrAlaProCys-----SerThrArgThrThrProThr 129
Qy 403 TCATCTATATGGCTCGCAACCCAGGATCTGGTGTCTTATATCAGTTCCACCGCT 462
Db 130 SerSer-----AlaThrProPro-----SerThrAlaThr 139
Qy 463 CTCTCGCGACCATGAGCTACCGGAGCACCTTTCAAGAATTCTGCCGAGGTTTATGAATG 522
Db 140 ValAlaArgProThrProProArgThrPro----- 149
Qy 523 ATAAAGCTGGGTACGGCTCCTGGTTGAGCAGCGTGCAGAGTTCTGGAGCACCGCATGG 582
Db 150 -----ProThrAlaValSerAlaGlyProArgGluTrp 160
Qy 583 ACTCGAAGCTGCTTTTCTCAAGTATGAACATGCATCGGAGCCTGGTGCAGTGGTGG 642
Db 161 SerProSerArg-----ArgSerTrp 167
Qy 643 AGCAGCTGGCCAGATTCTCGGGGTCTCTGTGACAAGGCCAGCTGGAAGCCCTGACGG 702
Db 168 ThrProTrpProAlaAlaTrp-----AlaArgIleArgTrpArgCysAlaSer 183
Qy 703 AGCACTGCCACAGCTGTGTG-----ACCAAGTGTGCAACGCTGAGGCC 747
Db 184 SerThrThrThrAlaArgThr-SerAlaThrSerProIleThrThrArgPro 200

RESULT 100
US-09-252-991A-30009
; Sequence 30009, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30009
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30009

Alignment Scores:
Pred. No.: 9,18e-07 Length: 195
Score: 159.50 Matches: 75
Percent Similarity: 36.7% Conservative: 15
Best Local Similarity: 30.6% Mismatches: 78
Query Match: 3.6% Indels: 77
DB: 2 Gaps: 15

US-10-768-158-1 (1-2419) x US-09-252-991A-30009 (1-195)
Qy 756 CCACGGCGAGCGGCTTCAG-----CGTTGCAGCACTGGTCCACCAGCTGGTGGCAGTGCT 703
Db 11 ProArgSerValProAlaAlaLeuArgCysArgAlaProProProAlaArgSerAlaGly 30
Qy 702 CCGTCAGGGCTTCAGCTGGGCTTGTTCACAGGACCCCGCAGGATCTGGCCAGCTGCT 643
Db 31 CysSerAlaCysThrArgGly---CysSerArg-----ProGlyCysTrpProAlaAla 47
Qy 642 CCACCATCGTCACAGGTCCTCGATCA-----TGTCTTCATACCTTGAGAAAAA 595
Db 48 ArgProValAlaProGlyAsnAspAlaSerAlaProGlyCysLeuArgArgAspGlyArg 67
Qy 594 GCACGTCGAGTCCATGCGGTGCTCCCGAGAACTCTCGACGCTGCTCAAAACGAGGACCGT 535
Db 68 SerArgArg-----ArgThrGlyCysArgArgArgThrAlaAlaAla 81
```

```
Qy 534 AGCCAGCTTATCATTAACCTCGCGAGAACTTCTGAAAGGTGCTCGGTAGTCA 475
Db 82 SerProProGlyGlySerArgArgTyrGlyArgGlyTyrGlyArgAlaArgGlyArgSer 101
Qy 474 TGGTCCGACAGAGCGGTGGAACCTGATAATAAGACACACCAGATCCT----- 427
Db 102 ArgArgProGlyArgAlaProArgProValArgLeuProAlaProGlyArgGlyPro 121
Qy 426 -----TGG-----GGTTGGAGCCATATAGATGACCTTGGAGTCTCCATTGGA 382
Db 122 GlySerSerArgTrpProGlyArgGluProAlaArg-----ArgSerValProSerGly 139
Qy 381 GGTCCAGAGGCGAGAAAGCGGTAGGGCAGGTGGCTCTTGTATGATGAGCGCGGAGAGTCA 322
Db 140 ProArgArgArgArgArgGlyArgGlyGly----- 150
Qy 321 CCTTGATGATGTCAGGCCCGGCTGTGGGTACTCCAGACCGGAGAGCTGCTCGTGTG 262
Db 151 -----GlySerProGly-----TyrArgSer 157
Qy 261 TCATCAAGCCGATCTCATCGGGTACCGCCCTGCTCACCAGTAGACACCACTCTCTGCA 202
Db 158 ThrAlaSerArgAsp-----GlyArgProAlaArgLeuProProAla 171
Qy 201 GCAAGCTGGTCCGGACTTGGGGTAGGTGACGATCCACACGTCGCTGGCGCCGACCGGGA 142
Db 172 -----Trp---ArgArgArgGlyArg----- 177
Qy 141 AGTTGGCGATCTCTCCATCTTCCCGCGGAGGAGGCGGCGCAGCCGACCATGGA 82
Db 178 -----ProProArgSerAspGly-----LeuSerHisLeuAla 188
Qy 81 CGAAGTACTGCTCT 67
Db 189 ThrAlaArgCysSer 193

Search completed: May 17, 2006, 10:01:45
Job time : 170.5 secs
```

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

☐ BLACK BORDERS

☒ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES

☐ FADED TEXT OR DRAWING

☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING

☐ SKEWED/SLANTED IMAGES

☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS

☐ GRAY SCALE DOCUMENTS

☐ LINES OR MARKS ON ORIGINAL DOCUMENT

☒ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY

☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

THIS PAGE BLANK (USPTO)